

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: Al Hansen Early
Searcher Phone #: 308-4501
Searcher Location: Biotech Lib
Date Searcher Picked Up: 8/24/01
Date Completed: 8/27/01
Searcher Prep & Review Time: _____
Clencal Prep Time: 4 min
Online Time: 2 min

Type of Search

NA Sequence (#) 2
AA Sequence (#) _____
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr.Link _____
Lexis/Nexis _____
Sequence Systems AB5501
WWW/Internet _____
Other (specify) _____

STIC-Biotech/ChemLib

49545

From: Jiang, Dong
Sent: Thursday, August 23, 2001 4:34 PM
To: STIC-Biotech/ChemLib
Subject: SN09/333,159

6/97
McCarthy, S.A.,
Fraser

Please search SEQ ID NO: 45 and 46

-issued
-commercial

Please send results on paper to Dong Jiang in 10B-01 (mail stop CM1-10C01). Thank you very much.

Dong Jiang (78243)
703-305-1345
U.S. Patent and Trademark Office
Art Unit 1646
dong.jiang@uspto.gov
CM1-10B01
Mail stop: CM1-10C01

Try to find the art only ≥ 54 nt (for 18 aa.)
| 60% var.

RECEIVED
AUG 24 2001
STIC

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:38:56 ; Search time 97.86 Seconds
(without alignments)
2454.893 Million cell updates/sec

Title: US-09-333-159-46

Perfect score: 1269
Sequence: 1 atgttgaaacctgtcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9455562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appl
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appl
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appl
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appl
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appl
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appl
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appl
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appl
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appl
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appl
11	46.4	3.7	43676	3 US-09-356-952-12	Sequence 12, Appl
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appl
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appl
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 154, App
15	36.6	2.9	3282	2 US-08-276-852-154	Sequence 154, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
17	36.6	2.9	3282	2 US-08-899-575-154	Sequence 154, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
19	36.6	2.9	3282	2 US-08-899-575-154	Sequence 154, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
21	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
23	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
25	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
27	36.6	2.9	13254	1 US-08-899-575-156	Sequence 170, App

28 36.6 2.9 13254 5 PCT-US95-08743-156 Sequence 156, App
29 36.6 2.9 13254 5 PCT-US95-08743-170 Sequence 170, App
30 33.6 2.6 8392 1 US-08-060-255-6 Sequence 6, Appl
31 33.6 2.6 8392 3 US-08-465-713-6 Sequence 6, Appl
32 33.6 2.6 8392 5 PCT-US93-05857-6 Sequence 6, Appl
33 32.6 2.6 7218 1 US-08-232-463-14 Sequence 14, Appl
34 32 2.5 2197 1 US-08-233-005-3 Sequence 3, Appl
35 32 2.5 2197 1 US-08-428-943-3 Sequence 3, Appl
36 32 2.5 2197 3 US-09-016-649-3 Sequence 3, Appl
37 32 2.5 2197 5 PCT-US95-04858-3 Sequence 3, Appl
38 32 2.5 8342 3 US-08-545-860D-63 Sequence 63, Appl
39 32 2.5 8342 5 PCT-US94-04496-63 Sequence 63, Appl
40 31 2.4 3212 3 US-08-673-814-1 Sequence 1, Appl
41 30.4 2.4 1529 1 US-08-336-778-1 Sequence 1, Appl
42 30.2 2.4 4982 3 US-08-699-1038-1 Sequence 1, Appl
43 30.2 2.4 10079 2 US-08-476-866-20 Sequence 20, Appl
44 30 2.4 3639 2 US-08-737-524B-26 Sequence 26, Appl
45 30 2.4 5178 2 US-08-474-169-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-227-108-2
; Sequence 2, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Caps 0;

QY 125 accagaagcattcatgaattagtgaaatcacaacatcaagagctatccctgtgag 184
DB 26 ACCCTGAAGTGACCAATGAATAGTCAGATGATCACCCTACTGGGATACCCAGCTGAGG 85
QY 185 aatagaagtcgaactgaagatgggtatcccttctgtttaacagagattccctcgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGACGGTTATATCCTTGGATCGACAGAAATCCCTTATGGGA 145
QY 245 tagtgaaccttaagaagacagggttccaggcctgtgttactgcagcatggcctagt 304
DB 146 GGAATAATTCAGAAATATAGCCGGAGACCTGTGTGATTTTTCACACGGTTTGTCTCG 205
QY 305 gaggtgctagcaactgatttccaaactgcccacaatagcctgggttccatttggcag 364
DB 206 CATCAGCCACAACTGATCTCCAACTGCGGCAACAGCCTGGCCTTATCCTGCGCG 265
QY 365 atgctgttttgacgtgtggtgaggggagcagcaggggaaacgcctgtgtcgaatacaca 424
DB 266 ACGCCGGTACGAGCTGTGCTGGGAAACAGCAGGGGCAACACCTGGCCGAGGAGGAAATC 325
QY 425 agacactctccatagacacagatgattctgtgggttccagtttatgatgagatggctagt 484
DB 326 TGTACTACTGCGCGGACTCGCTCGAATCTGGGCTTTTCAGCTTTGACGAGATGCTTAAT 385
QY 485 ttgacttctgcagtgataaaactttatttgcagaaaaacggccagggaaagattctatt 544
DB 386 ATGACCTTCCCGGACCACTTGCATTTTGAAGAAACGGGACGAGCAAGTACACT 445
QY 545 atgtcggtctattcacagggcaccacatggcttattgcatcttccacacatccagagc 604
DB 446 ACCTTGCCCATTTCCAGGGGACCAACATTTGTTTCATGCGCTTTTCCACCAATCCCAAGC 505
QY 605 tggctcgaataacaaatgatttctgtcttagcaccatagccactgttaagcatgcaa 664
DB 506 TGGCGAAACGGATCAAAACCTTCTATGCTAGCTCCGCTTGCACCGTGAATGATACCG 565
QY 665 aagccccgggacaaatttctgtccagatgatgatacagggattgtttggca 724
DB 566 AAACCTTGTAACAACTCATGCTCGTCCCTTCGTTCTCTTCAAGCTTATATTGGAA 625
QY 725 aaaaagaattctgtatcagacagatttctcagacaactgttattaccttggcc 784
DB 626 ACAAAATATTCTACCCACACCACTTCTTGATCAATTTTCGCCACCGAGGTATGCTCC 685
QY 785 agtgattcttgatcagattgttagtaatatcatgttacttctgtgtgattcaacacca 844
DB 686 GCGAGACGGTGGATCTCTCTGAGCAACGCCCTGTTTATCATTTTGTGGATTGACACTA 745
QY 845 acaatgaacatgagccagcaagtgatatgtctgccacacttctgtgaaacatctg 904
DB 746 TGAACCTTGAACATGACTCGCTGGATGTGTATCTGTGCATATATCCAGCAGGAACATCG 805
QY 905 tgcataattctacactgagccagcagtgatgaattctgtgaaactccggcattgact 964
DB 806 TTGAGAACGCTGCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTCCAAAGTTTGTACT 865
QY 965 ggggagtgagacacaaatctggaataatgcaatacagccaactcctgttaaggtacagag 1024
DB 866 GGGGAAGCCAGTTTCAAGACATGATGCATATCATCAGAGCATGCTCCCTACTACAACC 925
QY 1025 tcagagatatgacgggtccctacagcaatgtgacagggagtcaggactggtttcaaatc 1084
DB 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTGTCTGGCCGACC 985
QY 1085 cagaagacgtgaaatgctctctgtgagtgaccaaactcattaccataagaatattc 1144
DB 986 CTCAGATCTTGCCTTTTCCAGCTCCCAATCTCATTTTACACAGGAAGATTC 1045
QY 1145 ctgaatgggtcagtgatttcaatctgtgggttggatgctctccacagtgatgacaatg 1204
DB 1046 CTCCTTACAATCACTTGGACTTATCTGGGCGCATGGATGCCCCCTCAAGCGGTTTACAATG 1105

QY 1205 aaatcatccatctgatg 1221
DB 1106 AANTGTTTCCATGATG 1122
RESULT 2
US-09-073-674-2
; Sequence 2, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2

Query Match 35.0%; Score 444.2; DB 2; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 accagaagcattcatgaattagtgaaatcacaacatcaagagctatccctgtgag 184
DB 26 ACCCTGAAGTGACCAATGAATAGTCAGATGATCACCCTACTGGGATACCCAGCTGAGG 85
QY 185 aatagaagtcgaactgaagatgggtatcccttctgtttaacagagattccctcgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGACGGTTATATCCTTGGATCGACAGAAATCCCTTATGGGA 145
QY 245 tagtgaaccttaagaagacagggttccaggcctgtgttactgcagcatggcctagt 304
DB 146 GGAATAATTCAGAAATATAGCCGGAGACCTGTGTGATTTTTCACACGGTTTGTCTCG 205
QY 305 gaggtgctagcaactgatttccaaactgcccacaatagcctgggttccatttggcag 364
DB 206 CATCAGCCACAACTGATCTCCAACTGCGGCAACAGCCTGGCCTTATCCTGCGCG 265
QY 365 atgctgttttgacgtgtggtgaggggagcagcaggggaaacgcctgtgtcgaatacaca 424

266	DB	AGCCGGGTACGACGTGGCTGGGGAACACAGGGGCAACACTGGCCGACGAGGAATC	325
425	QY	agacactccatagacacaaatgagttctggggttttcagttatgatagdatggtcaggt	484
326	DB	TGTACTACTCGCCGACCTCCGTAATCTGGGGCTTTTCAGCTTTGACGAGATGGCTAAAT	385
485	QY	tgaactctctgcagtgataaactttattttgagaacacggccgagaaaaagatatctatt	544
386	DB	ATGACCTTTCCCGCCACCACTTACCTTCATCTTTGAAGAAACGGGACAGCAAGCTACACT	445
545	QY	atgtcggtctattccagagccaccacccatgggctttatgtatttccaccatgccagagc	604
446	DB	AGTTTGGCCATTCCACAGGGCACACCATTGGTTTCATCGCCTTTTCACCACTATCCCAAGC	505
605	QY	tggctcagaaaaatcaaaatgtattttgctttgagcaccatgagcactgtttaagcatgcaa	664
506	DB	TGCGAAGCGGATCAAAACCTTCTATGCATTAGTCCCGTTGCCACCGTGAAGTACACCG	565
665	QY	aaagcccgggaccacaaatttttggctgcccagataatgatacaagggattgtttggca	724
566	DB	AAACCCCTGTTAAACAAACTCATGCTGCTCCCTTCGTTCTTCAAGCTTATATTGGAA	625
725	QY	aaaaaatttctgtatcagaccagatttctcagacaaactgttatttacctttgtggcc	784
626	DB	ACAAATAATTACCCACACCACTCTTTTGATCAATTTCTGCGCACCGAGGTATGCTCCC	685
785	QY	aggtgattcttgatcagatttggtagtaatacatgttacttctctgggtgggattccaaccca	844
686	DB	GGGAGACGGTGGATCTCCTTCGACAAACGCCCTGTTTATCATTTGTGGATTTGACACTA	745
845	QY	acaatatgaacatgagccgacgaagtgtatatgtgcccacactctctgctgggaacatctg	904
746	DB	TGAACCTGAACATAGTGCCTTGGATGTATCTGTACATAATATCCACGACGAAACATCGG	805
905	QY	tgcacaaatattctacactggagccagcgagtgaaattctgttgaactccggcatttgaact	964
806	DB	TTCAGAACCTGCTCCACTGGTCCAGGCTGTTAAGTCTGGGAAGTTTCCAGCTTTTGACT	865
965	QY	gggggagtgagaccacaaatctggaaaaatgcaatccagccaaactcctgtgaagtgacagag	1024
866	DB	GGGGAAGCCAGTTTCAGAACATGATGCATATCATCAGAGCATGCCCTCCCTACTACAACC	925
1025	QY	tcaagagatgacggtccctacagacaaatgtggacagaggtcaggaatggcctttcaaatc	1084
926	DB	TGACAGACATGATGTGCCAATTCGCAGGTGTGGAAAGTGGCAACGACGATCTGCTGGCCGACC	985
1085	QY	cagaacagctgaaatgctgctctgaggtgaccaactcatcatcaataagaatttc	1144
986	DB	CTCACGATGTTGACCTTTTGCTTTCCAGAGCTCCCAATCTCATTTACACAGGAAGATTC	1045
1145	QY	ctgaatgggctcacgtggatttcattctgggttttggatgctctccacgcgcatgtacaatg	1204
1046	DB	CTCCTTACAATCACTTGGACTTATCTGGGCCATGGATGCCCTTCACAGCGGTTTACATG	1105
1205	QY	aaatcatcaatctgatg	1221
1106	DB	AAATTGTTTCCATGATG	1122

RESULTS

RESULT 3
US-08-227-108-4

US-08-227-108-4 ; Sequence 4, Application US/08227108 ; Patent No. 5807726

; GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude
APPLICANT: Blanchard, Claire

APPLICANT: JUNIPER, JEAN-LOUIS

APPLICANT: JUNIER, JEAN-LOUIS
TITLE OF INVENTION: Recombinant Dog Gastric Lipase

;	TITLE OF INVENTION:	REC
:	NUMBER OF SEQUENCES:	31

; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennle & Edmonds

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 35.0%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0;

Qy	125	accagaagcattcatgaatattagtgaaatcatccaaatcaagctatccctgtgag	184
Db	29	ACC.TGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGATACCCAGCTGAG	88
Qy	185	aattatgaatgcgaactgaagatgggtatatactcttcgtttaacagagattcctcgaagcc	244
Db	89	AATATGAAGTTGTACCGAAGACGGTTATATCTTGGATCGACAAATTCCTTATGGGA	148
Qy	245	tagtgcgaacttaagaagacaggttccaggcctgtgttacttcgacatggcctagtgtg	304
Db	149	GGAAAAATTGAGAGATATAGCGGGAGACCTGTGCTTTTTGCAACACGGTTGCTCG	208
Qy	305	gagtgctagcaactggatttccaaacctgccaaacaatagcctgggcttattcttggcag	364
Db	209	CAT.TAGCCACAAACTGGATCTCCAACCTGCCCAACACACCTTGCCTTCATCCTGGCG	268
Qy	365	atgtgtgttttgacgtgtgattggggaaacgaggggaaacgccttggtctcgaaaaaca	424
Db	269	ACG.TCGGGTACGAGCTGTGCTGGGAACAGCAGGGGCAACACCTGGGCACGAGGAATC	328
Qy	425	aga-actctccatagaccaagatgagttctgggctttcagttatgatgagatgctcaggt	484
Db	329	TGTTACTTCTCCCGCATCCGTGGAATCTTGGGCTTTTTCAGCTTTGACGAGATGGCTAAAT	388
Qy	485	ttgacctctctgcagtgataaaacttattttgcagaaaacgggccagggaaaaagatctatt	544
Db	389	ATG.ACCTTCCGGCCACCATTTGACTTTCATCTTGAAGAAAACGGGACAGGACAACTACACT	448
Qy	545	atgtcggtctattcacagggccaccacatttggtcttatttgcattttccacactgccagagc	604
Db	449	ACGTTGGCCATTCCACGGGCACCACTATGGTTTCATCGGCTTTTCCACCAATCCCAAGC	508
Qy	605	tgggtcgaaaaatacaaatgtatttttcttttagcaccacctagccactgttaagcatgcgaa	664
Db	509	TGGT.GAAACGGATCAAAACCTTCTATGCATTAGCTCCGTTGCCACGGTGAAGTACACCG	568

Db 809 TTCAGACGCTCTCCACTGGTCCAGCGCTGTAAGTCTGGAGATTCCAGCTTTGACT 868
Qy 965 gggggagtgagacacaaatctggaatgcaatcagccaaactcctgtaagggtacagag 1024
Db 869 GGGGAAGCCAGTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAAC 928
Qy 1025 tcagagatatagcgtccctacacaaatgtgacagaggtcagagactggctttcaatc 1084
Db 929 TGACAGACATGCATGTGCCAATCAGTGTGGAACGGTGGCAACGACTTGTGCGCCGACC 988
Qy 1085 cagaagcgtgaaatgctctctgagtgaccacacacacacacacacacacacacacac 1144
Db 989 CTCACGATGTTGACCTTTGCTTTCCAGCTCCCAATCTCTATTTACCAAGGAGATTC 1048
Qy 1145 ctgaatgggtcagtgatttcattctgggtgggtgggtgggtgggtgggtgggtgggtgg 1204
Db 1049 CTCCTTACAATCACTTGGACTTTATCTGGCCATGGATGCCCTCAAGCGGTTTACAATG 1108
Qy 1205 aaatcatccatctgatg 1221
Db 1109 AAATGTTTCCATGATG 1125

RESULT 5

US-08-227-108-6
; Sequence 6, Application US/08227108
; Patent No. 5807726

; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,108

; FILING DATE: 03-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fanucci, Allan A.

; REGISTRATION NUMBER: 30,256

; REFERENCE/DOCKET NUMBER: 7620-033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-9090

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1146 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

US-08-227-108-6

Query Match 35.0%; Score 444.2; DB 1; Length 1146;

Best Local Similarity 62.8%; Pred. No. 5.6e-139;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 125 acccagaagcattcgaattatgtgaaatcatccaaatcgaaggctatccctgtgagg 184

Db 26 ACCCTGAAGTGACCATGAATATAGTCAGATGATCACTACTCTGGGGATACCCAGCTGAGG 85
Qy 185 aatatgaagtcgaactgaagatgggtatatcccttctgtttaacagagattcctcagagcc 244
Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTCTTGGGATCGACAGAAATTCCTTATGGGA 145
Qy 245 tagtgcacacaaagacagaggttccagggcctgtggttactcagacatggcctagctg 304
Db 146 GGAATAATTCAGAGAAATATAGGCGGAGACCTGTGTGATTTTTCACACACGGTTGCTCG 205
Qy 305 gaggtgctagcaactgatttccaaactgcccacaaatagcctgggttcttattcctggcag 364
Db 206 CATCAGCCACAACTGGATCTCCAACTGCGCCCAACACAGCGCTGGCTTCATCTCTGGCGG 265
Qy 365 atgtcgtttttgacgtgtggtgggaacagcagggggaacagcgtgtgtctcgaacacaca 424
Db 266 ACGCCGGGTAGCAGCTGTGGCTGGGAACACAGAGGGGAACACCTGGGCCAGGAGGAATC 325
Qy 425 agacactctccatagacacaaagatgagttctgtgggttttcagttatgagatggctaggt 484
Db 326 TGTACTACTCGCCCGACTCCGTCGAATTTCTGGGCTTTTACGCTTTGACGAGATGGCTAAAT 385
Qy 485 ttgaccttctcagtcagtaaaacttttttcagaaaaacggccaggaagatctatt 544
Db 386 ATGACCTTCCCGCCACCATTGACTTCATCTTGAAGAAAACGGGACAGACAACTACACT 445
Qy 545 atgtcgttattcacaggggaccacacacacacacacacacacacacacacacacacacac 604
Db 446 ACGTTGGCCATTCCAGGGCACCACTTGTGTTTCATCGCCTTTTCCACCAATCCCAAGC 505
Qy 605 tggctcagaaaaatacaaatgtatttctttagcaccacacacacacacacacacacacacac 664
Db 506 TGGGAAACAGCGATCAAAACCTTCTATGCAATAGTCCCGTTGCCACCGTGAAGTACACGG 565
Qy 665 aaagcccggggacacaaattttgtgtcgcagatatgatgataagggatgtttggca 724
Db 566 AAACCTGTGTAACAACTCATGTCTGCTCCCTCTCTTCAAGCTTATATTTGGAA 625
Qy 725 aaaaagaattctgtatcagaccagatttctcagacaaactgttatttaccctttgtgccc 784
Db 626 ACATAATATTCTACCCACACCACATTCCTTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 685
Qy 785 aggtgattcttgatcagatttgtagtaataatcagtttacttctgtgggtgagttcaacaca 844
Db 686 GCGAGACGGTGGATCTCTCTGACAGAACGGCCTCTTTATCATTTTGTGGATTTGACACTA 745
Qy 845 acatataacatgagcagcagcaagtgtatgtgcccacacacacacacacacacacacacac 904
Db 746 TGAACCTGAACATGATGCTGGATGTGTCTCTCATATATCCAGCAGGAACATCGG 805
Qy 905 tgcacaaatattctacactggagccagggcagtgaaattctgtgaaactccgggctatttgact 964
Db 806 TTCAGAACGCTGCTCCACTGGTCCCGAGGCTGTTAAGTCTGGGAAGTTCGAAGCTTTTGACT 865
Qy 965 gggggagtgagacacacaaatctggaataatgcaatcagccaaactcctgtaagggtacagag 1024
Db 866 GGGGAAGCCAGTTCAGAACATGATGCACTATCATCAGACGATGCTCCCTACTACAAC 925
Qy 1025 tcagagatatgacgtccctacacaaatgtgacagaggttcagagactggctttcaaatc 1084
Db 926 TGACAGACATGCATGTGCCAATCAGCAGTGTGGAACGGTGGCAACGACTTGTGCGCCGACC 985
Qy 1085 cagaagcgtgaaatgctctcctgtgaggtgacacacacacacacacacacacacacacac 1144
Db 986 CTCACGATGTTGACCTTTTGTCTTCCAAAGCTCCCAATCTCATTTTACCAAGGAGATTC 1045
Qy 1145 ctgaatgggtcagtgatttctcctgtgggttttggatgctcctcagcagtatgtacaatg 1204
Db 1046 CTCCTTACAATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAAGCGGTTTACAATG 1105
Qy 1205 aaatcatccatctgatg 1221

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227.108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-227-108-1

Query Match 35.08; Score 444.2; DB 1; Length 1528;
Best Local Similarity 62.88; Pred. No. 6.8e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 accagaagcattcatgaatagtgaaatcattcaacacatcaaggctatccctgtgag 184
DB 26 ACCCTGAAGTACCAATGAATATAGTCAGATGATCACCCTACTGGGGATACCACCTGAGG 85

QY 185 aatagaagtcgaactgaagtggtatattccttctgttaacagagattcctcgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTTGGGATCGACAGATTCTTATGGGA 145

QY 245 tagtgaacacctaagaagcaggttccagcctgtgtgtgttactcagcagatggcctagttg 304
DB 146 GGAAAAATTCAGAGAAATATAGCGCGGAGACCTGTTGCATTTTTCACACACCGTTGCTCG 205

QY 305 gagggtcagcaactggtattcccaactgcccacaataagctgggtcttattctggcag 364
DB 206 CATCAGCCACAAATCTGGATCTCCAACTTGCACCAACAGCGCTGGCTTCTATCTGGCGG 265

QY 365 atcgtggtttgacgtgtggtggaacagcagcgagggaacgcctggtcttcgaaacaca 424
DB 266 ACGCCGGGTACGACGTGTGGTGGGAAACAGCAGCGGGCAACACCTGGGCCAGGAGATC 325

QY 425 agacactctccatagacacaaagatgagttctgtggtttcagttatgatgagatggctaggt 484
DB 326 TGTACTACTCGCCGACCTCCGCGAATCTGGGCTTTTCAGCTTTTCAGCAGATGGCTAAAT 385

QY 485 ttgactctcctcagtgataaacttatttgcagaaacggcgagcagaaagattctatt 544
DB 386 ATGACCTTCCCGCCACCAATGACTTCACTCTTGAAGAAACGGGACGACGACGACTACACT 445

QY 545 atgtcggtatttcacagggcaccaccatgggtctttattgcatctttccaccatgcagagc 604
DB 446 ACGTTGGCCATTCCCGAGGACCAACCATTTGTTTCATCGCTTTTCCACCAATCCCAAGC 505

QY 605 ttgctcagaataataatgtattttgttttagcaccatagcactgttaagaatgcaa 664
DB 506 TGGCGAAACGGATCAAAACCTTCTATGATGATAGCTCCGCTTTCGCCACCGTGAATACCG 565

QY 665 aaagccggggccaaattttgtgtgcgcagatgatgatgaagattgtttgca 724
DB 566 AAACCTGTAAACAACTCATGCTCGGCCCTTCTCTCAAGCTTATATTGGAA 625

QY 725 aaaaagaattctgatcagaccagattctcagacaactgtttattaccttttgcc 784
DB 626 ACATAATATCTACCCACCACTCTTTGATCAATTTCTGCCACCGGATGCTCCC 685

QY 785 agctgattcttgatcagattgttagtaataatcatgtttacttcttggttgattcaacacca 844
DB 686 GCGAGACGGTGGATCTCTCTGACAGCAACGCCCTGTTTATCATTTGTGGATTGACACTA 745

QY 845 acaatatgaacatgagccgagcaagtgatgtatgtgccacactctgtggaacatctg 904
DB 746 TGPACTTGAACATGAGTCGCTTGGATGTGTATCTGCATATATCCAGCAGGAACATCGG 805

QY 905 tgcaaaatattctacactgagccagcagtgaaattctgttgaaactccggcatttgact 964
DB 806 TTCAGAACGTCCTCCACTGGTCCAGGCTGTTAAGTCTGGAAGTCCCAAGCTTTTGACT 865

QY 965 gggggagtgagacaaaatctggaaaaatgcaatacagcacaactcctgtaaaggtaacagag 1024
DB 866 GGGGAAGCCCGAGTTCAGAACATGATGCACTATCATAGAGCATGCCCTCCCTACTACAACC 925

QY 1025 tcagagatatgacggtccctcacagcaaatgtgacagaggtcaggaactggctttcaaatc 1084
DB 926 TGACAGACATGCATGTGCCAATCCAGTGTGGACGGTGGCAAGACTTGTGTGGCCGACC 985

QY 1085 cagaagacgtgaaatgtcgtctctgtgagtgaccaacctcatctaccataagaatattc 1144
DB 986 CTCACGATGTTGACCTTTTGTCTTCCAAAGCTCCCAATCTCATTTACCACAGGAAGATTC 1045

QY 1145 ctgaatgggtcagctgaggtatttcattcctgttggttggtggtcctcaccgtatgtacaatg 1204
DB 1046 CTCCTTAAATCACTTGGACTTTATCTGGCCATGGATGCCCTCAAGCGGTTTACAATG 1105

QY 1205 aaatcattccatctgatg 1221
DB 1106 AAATGTTCCATGATG 1122

RESULT 8
US-09-073-674-1
; Sequence 1, Application US/09073674
; Patent No. 3998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073.674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cdna to mRNA
US-09-073-674-1

Query Match 35.0%; Score 444.2; DB 2; Length 1528;
Best Local Similarity 62.8%; Pred. No. 6.8e-139; Mismatches 0; Indels 0; Gaps 0;
Matches 689; Conservative 0;

QY 125 accagaagcattcatgaattatttagtgaatcatccaaacatcaagagctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGATACCCAGCTGAGG 85

QY 185 aatatgaagtgcgaactgaagtggtatatactctctgttaacagagattcctcagggcc 244
DB 86 AATATGAAGTTGTGACCAAGAGCGTTATATCTGGGATCGACAGATTCCTTATGGGA 145

QY 245 tagtgaacctgaagaagacaggttccagcctgtgtgttactgcagcagcttagtg 304
DB 146 GGAATAATTCAGAGATATAGCCGGAGACCTCTTGCAATTTTGCACACAGGTTTCTCG 205

QY 305 gaggtgtgacactgatttccaaactgtccaaacataagcctgggtcttattctggcag 364
DB 206 CATCAGCCACAACTGGATCTCAACTGCCCAACACAGCTGGCTTCATCTCTGGCCG 265

QY 365 atgctggttttgacgtgtggtgaggggagacagcaggggaaacgcctggtctgaaacaca 424
DB 266 ACGCCGGGTACGACGTGTGCTGGGAACAGCAGGGGCAACACCTGGCGCAGGAGAAATC 325

QY 425 agacactctccatagaccagaatgatttctgggtttcagttatgatgagtggttaggt 484
DB 326 TGACTACTGCGCCGACCTCGTGAATTTGGGCTTTTACGCTTTGACGAGATGGCTAAAT 385

QY 485 ttgacctctctgcagtgaataactttatttgcagaaacagggccaggaagactatt 544
DB 386 ATGACCTTCCCGCCACCATGTACTTCTTGAGAAACAGGGACAGACAGCTACAT 445

QY 545 atgctggttattcaggggacacacacatgggtttatttgcattttccaccatggcagagc 604
DB 446 ACCTTGGCCATTTCCAGGGCACCACCATTTGTTTCACTGCTTTTCCACCAATCCCAAGC 505

QY 605 tggctcagaaatcaaaatgtatttgccttttagcaccatagccactgttaagcatgcaa 664
DB 506 TGGCGAAACGGATCAAAACCTTCTATGATTAAGTCTCCCTGCGACCGTGAAGTACACCG 565

QY 665 aaagcccccgggaccataattttgtgctccagatatgatcatcaagggtattgttgca 724
DB 566 AAACCTGTTAAACAACTCATGCTCGTCCCTCTCTCTCAAGCTTATTTGGAA 825

QY 725 aaaaaaattctgtatcagaccagatttctcagacaactgttatttaccctttgtggcc 784
DB 626 ACAAATATTCTACCCACACCACTCTTTGATCAATTTCTCGCCACCGAGGTATGCTCC 685

QY 785 aggtgattctgacagattttagtaataatcatgttaattcttctggtgattcaacacca 844
DB 686 GCAGAGCGGTGATCTCCCTGTCAGCAACGCCCTGTTTATCATTTTGTGGATTGACACTA 745

QY 845 acaatatgaacatgagccagcaagtggtatgtctgcccacactcttctggaacatctg 904
DB 746 TGAACTTGAACATGATCTCGCTTGGATGTGTACTGTCTACATATCCAGCAGGAACATCGG 805

QY 905 tgcataattctacactggagccagggcagtggaattctggtgaactccgggcatcttgaact 964
DB 806 TTCAGAACGCTGCTCCACTGTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCCAAGCTTTGACT 865

QY 965 gggggagtgagaccacaaatctggaataatgcaatcagcacaactcctgttaaggtacagag 1024
DB 866 GGGGAAGCCCGATGTCAGACATGATGCATATCATCAGAGCATGCTCCCTACTACACCC 925

QY 1025 tcagagatatgacgttccctacagcattgtggacagaggtctcaggaactggcttcaaatc 1084
DB 926 TGACAGACATGCATGTGCCAATCGCAGTGTGNAACGGTGGCAACGACTTGTCTGGCCGACC 985

QY 1085 cagaagacgtgaaatgctgctctctgaggtgaccacactcatctaccataagaatattc 1144
DB 986 CTCAGGATGTTGACCTTTTGTCTTCCCAAGCTCCCAATCTCATTTACCACAGGAATTC 1045

QY 1145 ctgaatgggctcactggtgatttcactctctgggttttgatgctcctaccgtatgacaatg 1204
DB 1046 CTCCTTACAATCACTTGGAGTTTATCTGGGCCATGATGCCCTCAAGCGGTTTACAATG 1105

QY 1205 aaatcatccatctgatg 1221
DB 1106 AAATGCTTTCCATGATG 1122

RESULT 9
US-08-751-782-2
; Sequence 2, Application US/08751782
; Patent No. 5821352
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Johnathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A cDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 18-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.2
US-08-751-782-2

Query Match 3.8%; Score 48.2; DB 1; Length 178;
Best Local Similarity 55.8%; Pred. No. 1.4e-06;
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 416 gaaacacacagacactctccatccatagaccagaagatgagttctgggtttcagttatgaga 475
DB 8 GGAACATGTGGCCCTAGACCCAGATTCCTAAAGAAATTTGGGATTTAGTTTAAATGAAC 67

QY 476 tggctaggtttacccttctcctgcagtgataaaacttttttgcagaaaaacggccagaaa 535
DB 68 AAATAGATACGACCTCCCGCCCATCATTTATTTCATCTGTAATGAACAAGAACAC 127

RESULT 13
 US-09-163-669-1
 ; Sequence 1, Application US/09163669
 ; Patent No. 6111076
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: FUKUSUMI, SHOJI
 ;
 ; APPLICANT: HINUMA, SHOJI
 ;
 ; APPLICANT: FUJII, RYO
 ;
 ; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
 ;
 ; TITLE OF INVENTION: RECEPTOR (HIBCD07)
 ;
 ; NUMBER OF SEQUENCES: 14
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSEE: Rafter & Prestlia
 ;
 ; STREET: P.O. Box 980
 ;
 ; CITY: Valley Forge
 ;

RECORD 14
 US-08-276-852-154
 ; Sequence 154, Application US/08276852
 ; Patent No. 5652138
 ; GENERAL INFORMATION:
 ; APPLICANT: Burton, Dennis R
 ; APPLICANT: Barbas, Carlos F
 ; APPLICANT: Lerner, Richard A
 ; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
 ; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
 ; NUMBER OF SEQUENCES: 170
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: The Scripps Research Institute, Office of
 ; ADDRESSEE: Patent Counsel
 ; STREET: 10666 No. 552138th Torrey Pines Road, Suite 220,
 ; STREET: Mail Drop TPC8
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 154:
SEQUENCE CHARACTERISTICS:
LENGTH: 3282 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 15...452
US-08-276-852-154

Query Match 2.9%; Score 36.6; DB 1; Length 3282;
Best Local Similarity 51.5%; Pred. No. 0.074;
Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 435 catagaccaagatgagttctggttttcagttatgatgagatggttaggttgaccttc 494
Db 302 CACAGCCTACATGAGTGGAGGAGCTCAGGTGTCACACACGCGTGTATTATTGTCG 361
Qy 495 tgcagtgataaactttattttgcagaaaaacggccaggaagaatctattatgctggcta 554
Db 362 GAGAGTGGGGCCATATAGTTGGGATGATTCCTCCAGGACAAATTATTATATGGAGCTG 421
Qy 555 ttccaggccaccaccatggcctttattgcatcttccaccatg 597
Db 422 GGGCAAGGAACCGGTCATCGTGAGCTCAGCTTCCACCAAG 464

RESULT 15
US-08-276-852-169/c
Sequence 169 Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 3282 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-276-852-169

Query Match 2.9%; Score 36.6; DB 1; Length 3282;
Best Local Similarity 51.5%; Pred. No. 0.074;
Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 435 catagaccaagatgagttctggttttcagttatgatgagatggttaggttgaccttc 494
Db 2981 CACAGCCTACATGAGTGGAGGAGCTCAGGTGTCACACACGCGTGTATTATTGTCG 2922
Qy 495 tgcagtgataaactttattttgcagaaaaacggccaggaagaatctattatgctggcta 554
Db 2921 GAGAGTGGGGCCATATAGTTGGGATGATTCCTCCAGGACAAATTATTATATGGAGCTG 2862
Qy 555 ttccaggccaccaccatggcctttattgcatcttccaccatg 597
Db 2861 GGGCAAGGAACCGGTCATCGTGAGCTCAGCTTCCACCAAG 2819

Search completed: August 24, 2001, 22:39:52
Job time: 7308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:44:11 ; Search time 4395.77 Seconds
(without alignments)
4465.332 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgtcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v11: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_rol2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	DB	ID	Description
1	527.8	41.6	2481	93	HSUYACLY	231690	H.sapiens (
2	527.8	41.6	2481	97	HSU08464	U08464	Human lysos
3	527.8	41.6	2626	93	HSIAL	X76488	H.sapiens m
4	526.2	41.5	2493	97	HUMLIPOHL	M74775	Human lysos
5	489.2	38.6	1378	9	A26689	A26689	Precursor o
6	489.2	38.6	1378	9	A26690	A26690	Precursor o
7	470	37.0	1365	93	HSGLR	X05997	Human mRNA
8	470	37.0	1367	9	A01046	A01046	H.sapiens m

```
9 470 37.0 1367 9 A12714 Pregastric
10 467.6 36.8 1336 9 A01157
11 467.6 36.8 1355 95 RNLIP
12 461 36.3 2358 94 MMLYACLY
13 459.8 36.2 3144 95 S81497
14 458 36.1 1417 7 BOWPE
15 444.2 35.0 1137 9 AR039022
16 444.2 35.0 1137 9 AR092633
17 444.2 35.0 1140 9 AR039023
18 444.2 35.0 1140 9 AR092634
19 444.2 35.0 1146 9 AR039024
20 444.2 35.0 1146 9 AR092635
21 444.2 35.0 1198 9 A57760
22 444.2 35.0 1528 9 A57756
23 444.2 35.0 1528 9 AR039021
24 444.2 35.0 1528 9 AR092632
25 444.2 35.0 1651 7 CEFIPASE
26 441 34.8 1137 9 A39301
27 441 34.8 1140 9 A39303
28 441 34.8 1146 9 A39305
29 441 34.8 1528 9 A39305
30 388.8 30.6 1048 9 A57758
31 286 22.5 699 10 AX068249
32 286 22.5 1308 10 AX068259
33 269 21.2 76653 63 AC015506
34 269 21.2 158195 63 AC015461
35 269 21.2 181170 79 AL353113
36 204.4 16.1 148720 65 AC019353
37 204.4 16.1 172834 62 AC011277
38 151.2 11.9 76653 63 AC015506
39 147 11.6 161826 90 AL358532
40 140.6 11.1 401 93 HSLIPCHL04
41 140.6 11.1 1851 93 HSLIPAP4
42 140.6 11.1 180366 81 AL513533
43 127.6 10.1 40114 6 CELZK6
44 122.4 9.6 137911 61 AC009769
45 122.4 9.6 204102 63 AC013826
```

AL12714 Pregastric
A01157 R.norvegicus
X02309 Rat mRNA fo
Z31689 M.musculus
S81497 lysosomal a
L26319 Bos taurus
AR039022 Sequence
AR092633 Sequence
AR039023 Sequence
AR092634 Sequence
AR039024 Sequence
AR092635 Sequence
A57760 Sequence 5
A57756 Sequence 1
AR039021 Sequence
AR092632 Sequence
Y13899 Canis famill.
A39301 Sequence 2
A39303 Sequence 4
A39305 Sequence 6
A39300 Sequence 1
A57758 Sequence 3
AX068249 Sequence
AX068259 Sequence
AC015506 Homo sapi
AC015461 Homo sapi
AL353113 Homo sapi
AC019353 Homo sapi
AC011277 Homo sapi
AC015506 Homo sapi
U04288 Human lysos
X75491 H.sapiens L
AL513533 Homo sapi
AF067942 Ctenorhab
AC009769 Homo sapi
AC013826 Homo sapi

sig_peptide 34..96
gene /gene="LAL"
CDS 34..1233
34..1233
/gene="LAL"
34..1233
/gene="LAL"
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="CAA83495.1"
/db_xref="GI:506431"
/db_xref="SPTREMBL:O16529"
/translation="MKMREFLGLVCLVLTWLTLSHSEGRGKLTAVDPTNMNVSIIISYWG
GPSEYLVETEDGYILCNRIPIHGRKNHSDKGPVVFQHLGLADSNVNTNLANS
SIGFTLADAGFVDMGNSGNTWSRKHKTLSVQDFWAFSYDEMAKYDLPASINFI
NKTGOVYVGVHSGOCTTIGTAFISOIPELAKRIKMFALGPVAVFCTSPMAKLR
LPDLIKDLFGDKELPOSALFKNLGTHVTHVILKELGNCFLCFLGSENNLNMSR
VDVYTHSPAGTSVQNMHWSOAVAFQKFOAFDWGSSAKNYPHYNSOYPTTYNVKDL
VPTAVWSGGHDWLADVDYVNIILLTQITNVLFHESIPWEHLDIFWGLDAPWRLYNKII
NLMRKIQ"

variation 79
/gene="LAL"
/citation={1}
/replace="c"
97..1230
/gene="LAL"
/product="lysosomal acid lipase"
100
/gene="LAL"
/citation={1}
/replace="g"
2459..2464
polyA_signal 2480
polyA_site 692 a 477 c 513 g 799 t
BASE COUNT
ORIGIN

Query Match 41.6%; Score 527.8; DB 93; Length 2481;
Best Local Similarity 65.4%; Pred. No. 4.5e-146;
Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1;

ALIGNMENTS

```
RESULT 1
HSLYACLY 2481 bp mRNA PRI 17-FEB-1997
LOCUS H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
ACCESSION 231690
VERSION 231690.1 GI:506430
KEYWORDS LAL; lipase; lysosomal acid lipase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2481)
Du,H. and Gregory,G.A.
Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
Unpublished
2 (bases 1 to 2481)
Du,H.
Direct Submission
Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
OH, 45229-3039, USA
FEATURES
source
1..2481
Location/Qualifiers
/organism="Homo sapiens"
/strain="HepG2"
/db_xref="taxon:9606"
/clone="clone hLAL"
/tissue_type="liver"
/cell_type="hepatoma"
/clone_lib="human liver HepG2 cDNA library/lambda 2AP"
```


RESULT
A26689
5

QY 845 acaatataacatgagccgagcaagtgatgtatgtgtccacactcttctgctggaacatctg 904
 DB 856 CAAACTTGAACATGAGTGGCTGGTGTACGTGTACATATAATCCGACAGAACTTCAG 915
 QY 905 tgcataattctacactgagccagcagcagtggaattctgtggaactccgggcatcttact 964
 DB 916 TTAACAAATGCTGCACGTGGACCCAGGCTGTTAAATCTGGAATCTGGAATCTTCAAGCTTTAAAT 975
 QY 965 gggggagtgagacaaatctggaataatgcaatcagcccaactcctgttaagggtacagag 1024
 DB 976 GGGGAAGTCCAGCTCAGAAATGAGTGCACCTTCAATCAGCCACACCTCCCTACTACAAATG 1035
 QY 1025 tcagagatatagcaggtccctcagcaaatgttgagcagaggtcaggaactggtttccaaatc 1084
 DB 1036 TGACCGCATGAATGTGCAATTCAGTGTGGAGTGGTGGCAATGACTGGTGGCTGACC 1095
 QY 1085 cagaagcgtgaaatgctctcctgaggtgagcgaacacacctctatccataagaatattc 1144
 DB 1096 CCAAGATGTTGACCTTTTGCTTCCAAACTTTCTTAATCTCAATTTACCACAGGAGATTC 1155
 QY 1145 ctgaatgggctcagctgattctcctggtgttggatgctcctcaccgtatgtacaatg 1204
 DB 1156 TTCCATATAATCACTTGGATTTATCTGGGCAATGAATGCTTCCACAGAGTTTACAAATG 1215
 QY 1205 aaatccatctctgatgcagcagagggaga 1234
 DB 1216 AAATTTATTTCTATGATGCGCAAAAGATAAAA 1245

RESULT 6

A26690/c 1378 bp DNA PAT 05-APR-1995
 DEFINITION Precursor of rabbit gastric lipase (comp.).
 ACCESSION A26690
 VERSION A26690.1 GI:905030
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM Benicourt, C., Blanchard, C. and Junien, J.L.
 REFERENCE 1 (bases 1 to 1378)
 AUTHORS Recombinant gastric lipase from rabbit and pharmaceutical
 TITLE compositions
 JOURNAL Patent: EP 0542629-A 10 19-MAY-1993;
 FEATURES INSTITUT DE RECHERCHE JOUVEINAL
 Location/Qualifiers
 1..1378
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 418 a 276 c 287 g 397 t
 ORIGIN

Query Match 38.68; Score 489.2; DB 9; Length 1378;
 Best Local Similarity 65.08; Pred. No. 1.4e-134;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 125 acccagaagcattcatgaattagtgaaatcatccaacatcaaggtctatccctgtgagg 184
 DB 1243 ACCCTGAAGTGANTAGTAAATATAGTCAGATGATTTCTCTACTGGGGATACCCCAAGTGAAA 1184
 QY 185 aatagaagtcgcaactgaagtggtatctctcttctgttaacagaggattcctcagagcc 244
 DB 1183 AATATGAAGTTGTGACTCAGATGGTTATATCTTGTAGTTAACAGATTTCTTATGGGA 1124
 QY 245 tagtgcacctaagaagcagaggttccagggcctgtggtgttactcagcagatgcttagttg 304
 DB 1123 AGAAAAATTCAGGGAAACAGAGCCAGAGACCCGCTGATTTTTCAGCATGTTGCTGCTG 1064
 QY 305 ggggtgctagcaactggatttccacctgcccacaataagcctgggtcttattctggcag 364
 DB 1063 CATGAGCTCAAACTGGATCTCCCAACCTGCCCCAACACAGCGCTGGCCTTTATCTGGCAG 1004

QY 365 atgtgtgttttgacgtgtgatatggggaacagcaggggaaacgcctgtctgcgaaaaacaca 424
 DB 1003 ATGTGTGTTATGTGTGTCTGGCTGGGAACACAGCAGAGAAATACCTGGTCCAGAGAAACT 944
 QY 425 agaacctctccatagacagaagatagttctgtgggttttcagttatgtatgagatgcttgggt 484
 DB 943 TATATCTATTACCACAGACTCCGTTGAATTTCTGGGCTTTTAGCTTTGATGAAATGGCTAAAT 884
 QY 485 ttgaccttctcgtcagtgataaaacttttttgcagaaaaacggccaggaagaagatcctatt 544
 DB 883 ATGACCTTCCAGCCACCAATTCAGCTTCATTTGTAAGGAAACTGGACAGAAAGCTGCAC 824
 QY 545 atgtcggctattcacaggggaccaccattgggctttatttgcatttttccaccatgcccagagc 604
 DB 823 ATGTTGTGTCATTCACGGGACACCACTTGGTTTTATTGGCTTTTCTACAAATCCCAAGC 764
 QY 605 tggctcagaataacaaatgtatttttcttagcaccatagccactgtttaagaatgcaa 664
 DB 763 TGGTGAAGAAGTAACAAACCTTTTATGCAATTAGCTCCAGTCCGCACTGTGAAGTATACAA 704
 QY 665 aaagcccccgggaccaaatttttctgctgcccagatatgatcatcaagggtattgttggca 724
 DB 703 AAAGCCTTGTAAACAAACTTAGGTTTATTCCTCCAACCATGTTCAAGATTTATTTGGTG 644
 QY 725 aanaagaatttctgtatcagaccagatttctcagacaaactgttatttaccctttgtgccc 784
 DB 643 ACAAAATATTTCTACCCACACAATTTCTTTGATCAATTTCTTGGCACTCAAGTGTGTTCCC 584
 QY 785 aggtgattctctgacagattttagttagtaataatcatgttacttctgtggtggttcaacacca 844
 DB 583 GTGAGACACTGAATGTAAATTTGCAGCAATGCTTATTTATCATTTGTGGCTTTGACAGCG 524
 QY 845 acatatagaatcagcagcagcaagtgatgtatgctgcccacactctgtcagaacatctg 904
 DB 523 CAAACTTGAACATGAGTGGCTTGGATGTGACGTGTACATTAATCCGCGCAGAACTTCAG 464
 QY 905 tgcataattctctacactgagccagcagtgaaattctgtggaactccgggcatcttggact 964
 DB 463 TTCAAAACATGCTGCACCTGGACCCAGGCTGTTAAATCTGGAATTTTCAAGCTTTTAAT 404
 QY 965 gggggagtgagacaaataatctggaataatgcaatcagccaaactcctgttaagggtacagag 1024
 DB 403 GGGGAAGTCCAGCTCAGAAATGAGTGCACCTTCAATCAGCCACACCTCCCTACTACAATG 344
 QY 1025 tcaagatatagcgtccctacagcaatgtgacagaggtcagagctgcttcaaatc 1084
 DB 343 TGACCGCCATGAAATGTGCCAATTTGAGTGTGGAGTGGCAATGACTGGTGGCTGACC 284
 QY 1085 cagaagcgtgaaatgctcctcctgaggtgaccacacacctctatccataagaatattc 1144
 DB 283 CCAAGATGTTGACCTTTTGCTTCCAAACTTTCTTAATCTCATTTACCACAGGAGATTC 224
 QY 1145 ctgaatgggctcagtggtgattcctcctggtgttggatgctcctcagcagatgtgacaatg 1204
 DB 223 TTCCATATAATCACTTGGATTTTATCTGGGCAATGAATGCTCCTCAAGAAAGTTTACAATG 164
 QY 1205 aaatccatctctgatgcagcagagggaga 1234
 DB 163 AAATTTATTTCTATGATGCGCAAAAGATAAAA 134

RESULT 7

HSGLR
 LOCUS HSGLR 1365 bp mRNA PRI 30-MAR-1995
 DEFINITION Human mRNA for gastric lipase.
 ACCESSION X05997
 VERSION X05997.1 GI:31771
 KEYWORDS gastric lipase; lipase.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

FEATURES
  source      Location/Qualifiers
  sig_peptide 1..1367
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
  CDS         47..103
               /note="gastric lipase"
               /codon_start=1
               /product="lipase"
               /protein_id="CAA00125.1"
               /db_xref="GI:344242"
               /db_xref="SWISS-PROT:P07098"
               /translation="MWLLTMSASLISVLGTHGLFGLKLPSPVETMNIISOMITWYGY
PNEEYEVVTDGYLEVNRIPYKKNNGTGORPVVFLQHGLLASATNWSLNPNSL
AFILADAGYDVMGNSRNTWARRNLYSPDSVEFWAFSDENAKYDLPATIDFIVKK
TGOKOLHYVGHSGGTIGTAFSTNPSLAKRIKTFYALAPVATVKYTSILNKLREVP
QSLFKFIFGDKIPYHNFDFQATEVCSEMLNLCNALFICGFDKSNFTSRLD
VYLSHNPACTSVQNMHFHTQAVKSGKFOAYDNGSPVQNMHYDQSPPYNYVTAMNVP
IAVWNGKDLLADPDQVGLLLPKLPNIYHKEIPYHNLDFIWMADPAQEVYNDIVSM
ISDDKK"
  mat_peptide 104..1240
               /note="gastric lipase"
               /product="lipase"
  BASE COUNT 397 a 284 c 269 g 417 t
  ORIGIN
  Query Match 37.0%; Score 470; DB 9; Length 1367;
  Best Local Similarity 64.2%; Pred. No. 7.3e-129;
  Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
  QY 119 ctgtggaccagaagcattcatatattagtgaaatcatcccaacatcaaggctatccct 178
  DB 123 CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCCAA 182
  QY 179 gtaaggatgaagtcgaactgaagatggatattctcttctgttaacagatccctc 238
  DB 183 ATGAAGATATGAAGTTGTGACTGAGATGGTTATATCTTGAAGTCAATAGAATTCCTT 242
  QY 239 gaggcctagtgcacctaagaagacaggttccagggcctgtggtgttactcagcagtgcc 298
  DB 243 ATGGGAAGAAAATTCAGGGAATACAGCCGACAGACCTGTGTGTTTTCAGCAGTGTT 302
  QY 299 tagttgaggtgtagcaactgatttccaaacctgcccacaaatagcctggggttcattc 358
  DB 303 TGCTTGATCAGCCACCAAACTGGATTTCACACCTGCCGAACACAGCAGCTTGCCTTATTC 362
  QY 359 tggcagatgctgtttgacgtgtgagtgagggaacagcagggaacacgctgtctcgaa 418
  DB 363 TGGCAGATGCTGGTTATGATGTGTGGCTGGCAACAGCAGAGAAACACCTGGGCCAGAA 422
  QY 419 aacacaagacactctccatagaccagaatgagttctgtgggtttcagttatgatgagatgg 478
  DB 423 GAACCTTGACTATTACAGATTCAGTTCAATTCCTGGGCTTCAGCTTTGATGAATGG 482
  QY 479 ctaggtttgaccttcctcgcagtgataaacctttattttgcagaaaaacggccagaaaaa 538
  DB 483 CTAATAATGACCTTCCAGCCACCAATCGACTTCATTGTGTAAGAAAACCTGGACAGAGCAGC 542
  QY 539 tctattatgtcgtctacacagggcacacattgggtcttattgattttccaccatgc 598
  DB 543 TACACTATGTGGCCATTCACGGGACACCACTTGTGTTTATGGCTTTTTCACCAATC 602
  QY 599 cagagctggctcagaataatcaaatgtattttgcttttagcaccatagccactgtttaagc 658
  DB 603 CCACCTTGGCTTAAAGAAATCAAACTTCTATGCTCTAGCTCTGTTGGCAGCTGTGAAGT 662
  QY 659 atgcaaaagcccgggaccacaaattttttgtgtgcagatgatgatcatcaagggttgg 718
  DB 663 ATACAAAAGCCTTATAAACAACTTAGATTGTGTTCTCAATCCCTCTCAAGTTTATAT 722
  QY 719 ttggcaaaaagaagaattctgtatcagacacagatttctcagacaaactgttattacccctt 778
  DB 723 TTGCTGACAAAATATTCTTACCCACACAACTTCTTTGATCAATTTCTTGTCTACTGAAGTGT 782
  QY 779 gtggccaggtgattcttctgtagcagattttagtaataatcatcttactctctgggtgattca 838
  DB 783 GCTCCCGTGTGAGATGCTGAATCTCTTTTGACAGCAATGCCTTATTATAATTTGTGGATTG 842
  QY 839 acaccacaataatgaacatgagccgagcaagtgatgtatagtgtgccacactcttgcgtgaa 898
  DB 843 ACAGTAAAGAACTTTAAACAGAGTCGCTTGTGATGTGTATCTATCATACATAATCCAGCAGAA 902
  QY 899 catctgtgcaaaatatctacactgagccagcagcagtggaattctgtggaactccgggcat 958
  DB 903 CTTCCTGTTCAAAACATGTTCCATTTGACCCAGGCTGTTAAGTCTGGGAAATTCCAAGCTT 962
  QY 959 ttgactgggggagtgagaccacaaatcttgaaaaatgcaatcagccacactcctctgaagg 1018
  DB 963 ATGACTGGGGAGCCAGTTTCAGAAATAGGATGCACATGATGATGATGATGATGATGATGATG 1022
  QY 1019 acagagtcagagatatgacgggtccctacagcaatgtggacagaggtcagagactggttt 1078
  DB 1023 ACAATGTGACAGCATGAATGTACCAATTCAGTGTGGAACGGTGGCAAGACCTGTTGG 1082
  QY 1079 caaatccagaagacgtgaaaatgctctctcaggtgacacacactcatcaccataaga 1138
  DB 1083 CTGACCCCAAGATGTGGCTTTTGTCTCCAAACTCCCAACTCTTATTATCCACCAAGG 1142
  QY 1139 atattcctgaatgggtcagtcagtgatttcatctgggtgttggatgctcctccacgtatgt 1198
  DB 1143 AGATTCTTTTACATCACTTGGACTTTATCTGGGCAATGGATGCCCTCCCAAGAGTTT 1202
  QY 1199 acaatgaatcatccatctgat 1220
  DB 1203 ACAATGACATTGTTTCTATGAT 1224
  RESULT 9
  AL2714 1367 bp DNA PAT 09-AUG-1994
  LOCUS Pregastric lipase.
  DEFINITION Pregastric lipase.
  ACCESSION AL2714
  VERSION AL2714.1 GI:579560
  KEYWORDS human.
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 1367)
  AUTHORS
  TITLE
  JOURNAL
  FEATURES
  source
    1..1367
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /codon_start=1
    /product="pregastric lipase"
    /protein_id="CAA01053.1"
    /db_xref="GI:579561"
    /db_xref="SWISS-PROT:P07098"
    /translation="MWLLTMSASLISVLGTHGLFGLKLPSPVETMNIISOMITWYGY
PNEEYEVVTDGYLEVNRIPYKKNNGTGORPVVFLQHGLLASATNWSLNPNSL
AFILADAGYDVMGNSRNTWARRNLYSPDSVEFWAFSDENAKYDLPATIDFIVKK
TGOKOLHYVGHSGGTIGTAFSTNPSLAKRIKTFYALAPVATVKYTSILNKLREVP
QSLFKFIFGDKIPYHNFDFQATEVCSEMLNLCNALFICGFDKSNFTSRLD
VYLSHNPACTSVQNMHFHTQAVKSGKFOAYDNGSPVQNMHYDQSPPYNYVTAMNVP
IAVWNGKDLLADPDQVGLLLPKLPNIYHKEIPYHNLDFIWMADPAQEVYNDIVSM
ISDDKK"
  CDS
    47..1243
    /product="gastric lipase"
    /protein_id="CAA01053.1"
    /db_xref="GI:579561"
    /db_xref="SWISS-PROT:P07098"
    /translation="MWLLTMSASLISVLGTHGLFGLKLPSPVETMNIISOMITWYGY
PNEEYEVVTDGYLEVNRIPYKKNNGTGORPVVFLQHGLLASATNWSLNPNSL
AFILADAGYDVMGNSRNTWARRNLYSPDSVEFWAFSDENAKYDLPATIDFIVKK
TGOKOLHYVGHSGGTIGTAFSTNPSLAKRIKTFYALAPVATVKYTSILNKLREVP
QSLFKFIFGDKIPYHNFDFQATEVCSEMLNLCNALFICGFDKSNFTSRLD
VYLSHNPACTSVQNMHFHTQAVKSGKFOAYDNGSPVQNMHYDQSPPYNYVTAMNVP
IAVWNGKDLLADPDQVGLLLPKLPNIYHKEIPYHNLDFIWMADPAQEVYNDIVSM
ISDDKK"
  BASE COUNT 397 a 284 c 269 g 417 t
  ORIGIN

```


QY 479 ctagggttgacattctcagtgatataaacttatttttgcagaaacacggccagagaaaga 538
 Db 440 CTAAATATGACCTTCCGCCCAATAAACTTCTATGTGTAAGAAATGAGCAAGAGAGA 499
 QY 539 tctattatgtcgtctattcacagggcaccacacagggctttatgacatttccaccatgc 598
 Db 500 TACACTATGTTGGTCATCTCAGGCACCACTATGTTGTTTATGCTTCTTACCAATC 559
 QY 599 cagagctggctcagaaaaatcattgtatttgcagccacacacacactgcttaagc 658
 Db 560 CTACACTGGCCCAAAAAATCAAGAGCTTTTATGATTAAGTCCAGTTCCTACCGTGAAGT 619
 QY 659 atgcaaaagcccgagacaaatttttgcgtgcagagatatgatcaagagattgt 718
 Db 620 ATACAAAGTCCCTTGAAAAAGATTTTCAATTTATTCCTACATTTCTTCTCAAGCTTATGT 679
 QY 719 ttggcaaaagaaatttctgtatcagacagagatttctcagacacactgtattacatt 778
 Db 680 TTGGCAGAAATGTTCTGCCCCACACCTACTTTGATGACTTCTTCTGGTACCGAAGTGT 739
 QY 779 gtggccaggtgattctgtcagatattgttagtaatatcatgttacttctgggtgattca 838
 Db 740 GCTACGGGAGGTTCTAGATCTCTCTGCAGCAACACTTATTCATCTCTGTGGATTG 799
 QY 839 acaccaacaatatgaacatgagcagcagcaagtgatatgctgcccacacactcttctgtaa 898
 Db 800 ACNAGAAAACTTAAATGTGAGTCGTTTGTGATGTATCTAGGGCAATAATCCAGCAGGA 859
 QY 899 catctgtcaaaaatttctacatgagcagcagcagtgaaattctggtgaactccggcat 958
 Db 860 CATCTGTTCAAGACTTCTCCACGTGGGACAGCTGTTGATGATCTGGGAAATTTCAAGCCT 919
 QY 959 ttgactggggagtgagacaaaatctggaaaaatgcaatcagcccaactctctgaaggt 1018
 Db 920 TCAACTGGGAGGCCATCCAGAACATGTTACACTACACAGCAAGAGCTCTCTGAAT 979
 QY 1019 acagagtcagagatatgacggttccctacagcaatgtgagcaggggctcagagactggttt 1078
 Db 980 ATGATGTCAGCCACTGCTGCGAGTGTGAGTGTGGAAGCGTGGCAATGACATCTCTGG 1039
 QY 1079 caaatccagagagcgtgaaatgctgctctgaggtgagcaactctctaccataaga 1138
 Db 1040 CTGATCCCCAGAGTGCCTGATGCTTCCCAACTCTCCAACTCTCTGTTCCATAGG 1099
 QY 1139 atattctgaatgggctcagctgatttctctctgggttgggtgctctcaccgtatgt 1198
 Db 1100 AGATTCTTGCCTACAACTCACCTGGACTTCATCTGGCAATGATGCTCCCTCAAGAGTTT 1159
 QY 1199 acaatgaatcatccatctgatgagcagggagagaccacc 1240
 Db 1160 ACAATGAGATGATTTCCATGATGTCAGAAAGACTAAAGAATC 1201

RESULT 11
 RNLIP RNLIP 1355 bp mRNA ROD 30-MAR-1995
 LOCUS Rat mRNA for lingual lipase.
 DEFINITION X02309
 ACCESSION X02309
 VERSION X02309.1 GI:56595
 KEYWORDS glycoprotein; lipase.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1355)
 REFERENCES
 AUTHORS Docherty A.J., Bodmer M.W., Angal S., Verger R., Riviere C.,
 Lowe P.A., Lyons A., Entage J.S. and Harris T.J.
 TITLE Molecular cloning and nucleotide sequence of rat lingual lipase
 cDNA
 JOURNAL Nucleic Acids Res. 13 (6), 1891-1903 (1985)
 MEDLINE 85215587

COMMENT Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.
 FEATURES
 source Location/Qualifiers
 1. 1355
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 sig_peptide 49..102
 /product="put. signal peptide (aa -18 to -1)"
 CDS 49..1236
 /note="lipase precursor"
 /codon_start=1
 /protein_id="CAA26179.1"
 /db_xref="GI:56596"
 /translation="MWLLITSVISTFGGAHLGKLGPGNPEANMNISOMITWGYG
 COYEVYEDGYILGVYRIPIHGNKNSNIGKRPVYVYLQHLGIASATNMNIALPNNSLA
 FMLADGYDVLGNSRGNTSRKKNVYSPDSFEMAFYDFDEMAKYDLPATNFIVQKT
 GOKIHYVHSGQTTIGFIAPSTNPLAKIKTFYALAPVATVKTOSPLAKISPIPT
 FLKLMGKEMFLPHYTFDDFLGTEVCSREVLDDLCSNTLIFCGFDDKKNLVSFPDV
 YLGNPAGTSVQDFLHWAQLVRSKGFANWGPSQNMHLNQNTPPTDYDSAMTPVP
 AVWNGNDILADPDQVAVMLLPKLSNLLFHKELIAYNLHLDIFINWMDAPQEVYNEMISM
 AED"
 mat_peptide 103..1233
 /product="lipase (aa 1-377)"
 misc_feature 505..507
 /note="aa 153 pot. lipid binding site"
 misc_feature 1355
 /note="polyA site"
 BASE COUNT 400 a 291 c 276 g 307 t 1 others
 ORIGIN
 Query Match 36.8%; Score 467.6; DB 95; Length 1355;
 Best Local Similarity 63.5%; Pred. No. 3.8e-128;
 Matches 713; Conservative 0; Mismatches 409; Indels 0; Gaps 0;
 QY 119 ctgaggacccagagcattcatgaatatagtgaaatcatccacacacaggtatccct 178
 Db 122 CTGAAACCCCTGGAAGAAATATGAATATTTAGTCAGATGATACTTACTGGGATATCCAT 181
 QY 179 gtgaggaaatgaagtcgaactgaagatgggtatattcttctgttaacagagattccc 238
 Db 182 GTCAGAAATATGAGTTGTTACTGNAGATGGCTACATTTCTGGGGGTCTACAGAAATCCCTC 241
 QY 239 gaggcctagtgcacacacacagagctccagcagctggtggttactgcagcagctgccc 298
 Db 242 ATGGGAAGAAATATCTGAAATATAGCAAGAGACCTGTGTGTATTTGACAGCATGTT 301
 QY 299 tagttgaggtgctagcaactggatttccacacacacacacacacacacacacacacacac 358
 Db 302 TGA?CGCATCAGCCACAACTGGATTGCAAAATCTACCAACAAACAGCCTGGCCTTTATGC 361
 QY 359 tggcagatgctggtttgacgtggtgagtgagggaacagcaggggaacacacacacacacac 418
 Db 362 TAGCAGACGCTGGCTGTATGTGTGGCTGGGGAACAGTCGAGGAAATACATGGTCCAGGA 421
 QY 419 aacacacacacactctccatagacacacagatgagttcttctggcttccagttatgatgagtg 478
 Db 422 AAATGTATATACTACTACACACTCAGTTGAATTTCTGGGCTTTTACGCTTTTGTAGAAATGG 481
 QY 479 ctgagtttgaccttctcagtcagtgataaaacttattttgcagaaaaacggccaggaaga 538
 Db 482 CTAAATATGACCTTCCGCCCAATAAACTTCAATTTATGTTGTTTCTTCTTCTTCTTCTTCT 541
 QY 539 tctattatgtcgttattcagggcaccacacacacacacacacacacacacacacacacacac 598
 Db 542 TACACTATGTTGGTCTATCTCAGGGCACCACTATGTTGTTTCTTCTTCTTCTTCTTCTTCT 601
 QY 599 cagagctggctcagaaaaatcattgtatttctttagcaccacacacacacacacacacacacac 658
 Db 602 CTACACTGGCCCAAAAAATCAAGAGCTTTTATGATTAAGTCCAGTTCCTACCGTGAAGT 661
 QY 659 atgcaaaagcccgagacaaatttttgcgtgcagagatatgatcaagagattgt 718

Db 1141 TCAAGCTTTTACGCTGGGGAACGACGTGACAAAGATTATTTTTCATTACAAATCAGAGTTATCC 1200
 Qy 1011 tgytaaggctacagagctcagagatgacggtccctacacagcaatgtggacagagagtgcaaga 1070
 Db 1201 ACCCTTGTACAGCATAAAGACATGCAACTACCCACCTGCTTTATGGAGCGGGGCAAGGA 1260
 Qy 1071 ctggctttcacaatccagaagacgctgaaatgctgctctctctgaggtgacccacccatccta 1130
 Db 1261 CTGGCTGGCAGACACACGAGTGACATCAATATCTTACTGACTGAATCCCCAGCTTAGTGTA 1320
 Qy 1131 ccataagaataattcctgaatgggtcagctcagtgagattcctcctgaggttgatgctcctca 1190
 Db 1321 CCACAAGAACATTCCTGGAGTGGGACCATCTGGATTTATCTGGGGTTTGGATGCCCTG 1380
 Qy 1191 ccgtatgtacataaataatccatcctcctgagtcagcagcagagag 1233
 Db 1381 GAGGCTGTATTAAGTAGTCACTCTANTGAAGAAATACCAAG 1423

RESULT 14

BOVME BOVME 1417 bp mRNA MAM 10-DEC-1994
 LOCUS Bos taurus pregastric esterase mRNA, complete cds.
 DEFINITION L26319
 ACCESSION L26319
 VERSION L26319.1 GI:600756
 KEYWORDS esterase; pregastric esterase.
 SOURCE Bos taurus 25 weeks Tongue serous gland cDNA to mRNA.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 1417)
 Timmermans,M.Y.J., Kupers,L.P. and Teuchy,H.
 TITLE The cDNA sequence encoding bovine pregastric esterase
 JOURNAL Gene 147, 259-262 (1994)
 MEDLINE 95011625
 FEATURES Location/Qualifiers
 1..1417
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /cell_type="Secretory cell"
 /dev_stage="25 weeks"
 /tissue_type="Tongue serous gland"
 73..129
 /EC_number="3.1.1.3"
 /codon_start=1
 /product="pregastric esterase"
 /protein_id="AA57037.1"
 /db_xref="GI:600757"
 /translation="MWWLLVTVCFIHMSGNAFCFLGKIKNPEASNNVSMISYWCYP
 SEMHKVITADGYILOVYRPHGNHNNHGORPVVFLHGLLGSATNWSNLPKNSLG
 FLADAGYDVLGNSRGNTWAOHLVYSPSPFWAFSEDEAEYDLPSTIDFILRPT
 GOKLHYVCHSOGTTGFIATFSTPLAEKIKVFYALAPVATVTKSLFNKLALIPH
 FLEKIFGDKMTPIHFLQFLGVEMCSRETLVLCNKALFAITGVNDKNFMSRLDY
 YIAHNPAGTSVONTLHWRAVSKGFQAFDQAPYQNLNMYHQPTPIYNLFAMNVP
 AYSADNDLLADPDQDFLLSLNLIYHKEIPNYNHLDFIWMADAPQEVYNEIVSLM
 AEDKK"
 130..1263
 /EC_number="3.1.1.3"
 /function="triacylglycerol acylhydrolase"
 /evidence="experimental"
 /product="pregastric esterase"
 1378..1385
 polyA_signal
 polyA_site 1396
 BASE COUNT 408 a 306 c 276 g 427 t
 ORIGIN

Query Match

Best Local Similarity 36.1%; Score 458; DB 7; Length 1417;
 Matches 741; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

Qy 44 tggaaatgtggcttctgattctggtggttatgtttccagagaaaatgtgaattcagttac 103
 Db 68 TGAATAATGTGGTGGCTACTTGTAAACAGTGTGTTTCATCCACATGCTCTGGAATGCAATTT 127
 Qy 104 atatgc---caactaaagctgtggccacagagcattcatgaataatagtgaaacatccc 160
 Db 128 GTTTCTTTGGAAAAATTCCTAAGAACCTGAAGCCAGTATGAATGTTAGTCAGATGATTT 187
 Qy 161 aacataagcctatccctgtgaggaatgatgaagtgcgaactgaagatgggttatcccttt 220
 Db 188 CTTACTGGGGCTACCCAAAGTGAATGAATAAGTTAACTGCGGATGTTTATATCCCTTC 247
 Qy 221 ctgttaacagagattcctcagagccctagtgaacacctgaagaacaggttccagggcctg 280
 Db 248 AGGCTATCGGATTCCTCATGGAAGAATAATGCTAATCATTTAGGTGAGAGACCTGTTG 307
 Qy 281 tgttactgcagcatggcctagtgtgaggtgctgtagcaactggatttccaaacctgcccaaca 340
 Db 308 TGTTCCTGCAGCATGGTCTTCTTGGATCAGCCACAACTGGATTTCCAACCTGCCAAGA 367
 Qy 341 atagcctgggcttctcctgtgagagatgctggtttgtgacgtgtgagatggggaaacagcagg 400
 Db 368 ACAGCTGGGCTTCTCTCTGCGCAGATGCTGTTATGACGTGTGGTGGGGAACAGCAGAG 427
 Qy 401 gaaacgctggtctcggaaacacacacacactctccatagacacagagatgcttgggctt 460
 Db 428 GAAACACCTGGGCGCCAGGAACATTTATCTATTCACAGACTCCCGGAATCTTGGGGCTT 487
 Qy 461 tcaagtatgatgagatggctaggtttgacacctctcctgctgagtataaactttattttgcaga 520
 Db 488 TCAGCTTTGTGTAATGCGGGAATATGACCTTCCATCTACAAATGATTTTCATCTAAGGA 547
 Qy 521 aaacgggccaaggaagaatctattatgctggtattcacaggggacacacacatgggcttta 580
 Db 548 GAACAGGACAGAAGAGCTACATATGTTGGCCATTGCCAAGGACCAACCATTTGGTTTA 607
 Qy 581 ttgcattttccacatgccagagcgtggtctcagaaaaatcctgataaaactttattgtcttagcac 640
 Db 608 TCGCTTTTCTACAGTCCCATGCTGGCTGAAAAAATCAAGTCTTCTATGCATTAGCCC 667
 Qy 641 ccatagccactgttaagcatgcaaaaagcccgaggacaaatattttgtgtgctgcagata 700
 Db 668 CAGTTGCCACAGTGAAGTACACCAAGAGCCTGTTTAAACAAACTTGCATTTATCTCCTACT 727
 Qy 701 tgatgatcaagagatgtttggcaaaaaaagtctgtatcagacagagatttctcagac 760
 Db 728 TCCTCTTCAAGATTATATTGGTGACAAAATGTTCTACCCACACACTTTTGTGAACAAT 787
 Qy 761 aactgttattaccctttgtggccaggtgattcttgcagagatttgcagtaataatcatgt 820
 Db 788 TTCTTGGTGTGAATGTGCTCCGTGAGACACTGGATGCTCTTGTGAAGATGCCCTGT 847
 Qy 821 tacttctgggtgattcaacacacacaaatatagaactagccgagcaagtgtatctgt 880
 Db 848 TTGCCATTACTGGAGTTGACATAAAAACTTCAACATGAGTGCCTTATAGTGTATATAG 907
 Qy 881 cccacactctgtggaacatctgtcaaaaatattctacactggagcagcagcagtaatt 940
 Db 908 CACATAATCCAGCAGGAACCTTCTGTTCAAAACACCCCTCCACTGGAGACAGCTCTTAAGT 967
 Qy 941 ctggtgaactccgggcatcttgactgggggagtgagacacaaaactctgaaaaatgcacatc 1000
 Db 968 CTGGGAAATTCCAAGCTTTTGACTGGGAGCCCCCATATCAGAACCTTAATGCATTATCATC 1027
 Qy 1001 agccaactcctgtaaggtacagagtcagagatatgacggttccctcacagcaatgtggacag 1060
 Db 1028 AGCCACACCTCCCATCTACAAATTTAACAGCCATGAATGTCCTCAATTTGCAGTAGTGAGTG 1087
 Qy 1061 gaggtcagactggcttccaaatccagagaagacgtgaaaaatgctgctctcaggtgacca 1120
 Db 1088 CTGACAATACCTGTGGGTGACCCCTCAGGATGTTGACTTTTCTGCTTTTCAAAACTCTCTA 1147

GenCore version 4.5
Copyright (C) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:41:42 ; Search time 211.8 Seconds
(Without alignments)
3762.073 Million cell updates/sec

Title: US-09-333-159-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtcaag.....gacggtgtgagccggtattg 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/NA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/NA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/NA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/NA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/NA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/NA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/NA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/NA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/NA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/NA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/NA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/NA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/NA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/NA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/NA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/NA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/NA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/NA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/NA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/NA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/NA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1269	100.0	1269	22	AAF45132
2	1269	100.0	2044	22	AAF45131
3	489.2	38.6	1378	14	AAQ42310
4	470	37.0	1367	7	AAAG0685
5	470	37.0	1367	7	AAAG0566
6	470	37.0	1367	17	AAAT5816
7	464.4	36.6	1366	6	AAAT50385
8	444.2	35.0	1528	15	AAQ68388
9	444.2	35.0	1528	17	AAAT5815
10	444.2	35.0	1531	15	AAQ68389
11	286	22.5	699	22	AAF28679

12	285	22.5	1308	22	AAF28689
13	257.3	20.3	617	21	AAA44349
14	193.2	15.2	572	21	AAA43744
15	100	7.9	1718	21	AAZ34958
16	89.4	7.0	1483	21	AAZ34956
17	84.8	6.7	859	21	AAZ34953
18	72	5.7	1438	21	AAZ34959
19	68	5.4	936	22	AAF58252
20	68	5.4	936	22	AAF58254
21	68	5.4	936	22	AAF58257
22	68	5.4	936	22	AAF58259
23	68	5.4	936	22	AAF58262
24	68	5.4	938	22	AAF58255
25	66.2	5.2	936	22	AAF58252
26	66.2	5.2	936	22	AAF58254
27	66.2	5.2	936	22	AAF58257
28	66.2	5.2	936	22	AAF58259
29	66.2	5.2	936	22	AAF58262
30	66.2	5.2	938	22	AAF58255
31	55.6	4.4	705	21	AAZ34952
32	54	4.3	549	21	AAC93827
33	51	4.0	509	21	AAZ34954
34	48.2	3.8	178	19	AAV35431
35	46.4	3.7	43676	21	AAZ60606
36	38.6	3.0	1664976	19	AAV21209
37	36.8	2.9	13254	21	AAA32165
38	36.6	2.9	244	22	AAF58238
39	36.6	2.9	3271	19	AAV00060
40	36.6	2.9	3282	17	AAT40914
41	36.6	2.9	3282	21	AAA31024
42	36.6	2.9	3282	21	AAA31038
43	36.6	2.9	3282	21	AAA32150
44	36.6	2.9	3282	21	AAA32164
45	36.6	2.9	5412	20	AAZ21356

ALIGNMENTS

RESULT 1

AAF45132
ID AAF45132 standard; cDNA; 1269 BP.

XX
AC AAF45132;

XX
DT 30-MAR-2001 (first entry)

XX
DE Human TANGO 294 ORF.

XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.

XX
OS Homo sapiens.

XX
PN WO200077239-A2.

XX
PD 21-DEC-2000.

XX
PF 24-MAY-2000; 2000WO-US14858.

XX
PR 14-JUN-1999; 99US-0333159.

XX
PA (MILL-) MILLENNIUM PHARM INC.

XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX
DR WPI; 2001-032313/04.

XX
DR P-PSDB; AAB66065.

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 XX disease.

XX Claim 1; Fig 6; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45131-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

Query Match 100.0%; Score 1269; DB 22; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgttggaacccctgtcaagacagtgattgtctcacacagaaatggaatgtggtctctg 60
 Db 1 atgttggaacccctgtcaagacagtgattgtctcacacagaaatggaatgtggtctctg 60
 Qy 61 attctgtgtggtatattgttccagagaaatgtgaattcagtagatgccaactaaagct 120
 Db 61 attctgtgtggtatattgttccagagaaatgtgaattcagtagatgccaactaaagct 120
 Qy 121 gtgacacacagaaagattcagaattattagtgaaatccacacatacgaagctatccctgt 180
 Db 121 gtgacacacagaaagattcagaattattagtgaaatccacacatacgaagctatccctgt 180
 Qy 181 gaggaatatgaatgcgcaactgaagatgggtatattcttctgtttaaagattcctcga 240
 Db 181 gaggaatatgaatgcgcaactgaagatgggtatattcttctgtttaaagattcctcga 240
 Qy 241 ggctagtgaacctaagaagacaggttccagcctgtgtgttactgcagcatggccta 300
 Db 241 ggctagtgaacctaagaagacaggttccagcctgtgtgttactgcagcatggccta 300
 Qy 301 gttgaggtgttagcaactggattcccaactgcacacataagctgggttccattctg 360
 Db 301 gttgaggtgttagcaactggattcccaactgcacacataagctgggttccattctg 360
 Qy 361 gcagatgctggtttgtgacgtgtggtgggaacagcggggaacgcctgtgtcgaata 420
 Db 361 gcagatgctggtttgtgacgtgtggtgggaacagcggggaacgcctgtgtcgaata 420
 Qy 421 cacaagacactctccatagaccagaagatgattcttggccttccagttatgatgatggct 480
 Db 421 cacaagacactctccatagaccagaagatgattcttggccttccagttatgatgatggct 480
 Qy 481 aggtttgaccttctcgtcagtgataaactttattttgcagaaaaacggccaggaagaatc 540
 Db 481 aggtttgaccttctcgtcagtgataaactttattttgcagaaaaacggccaggaagaatc 540
 Qy 541 tattatgctggttattcacagggcacaccatgggttttattgcatatttccaccatgcca 600
 Db 541 tattatgctggttattcacagggcacaccatgggttttattgcatatttccaccatgcca 600
 Qy 601 gagctgctcagaaaaacaaaatgtatttggcttagcaccatagccaactgttaagcat 660
 Db 601 gagctgctcagaaaaacaaaatgtatttggcttagcaccatagccaactgttaagcat 660
 Qy 661 gcaaaaagcccgaggcacaattttgtgtcctcagatgatgatcaaggatgtttt 720
 Db 661 gcaaaaagcccgaggcacaattttgtgtcctcagatgatgatcaaggatgtttt

Db 661 gcaaaaagcccgaggcacaattttgtgtcctcagatgatgatcaaggatgtttt 720
 Qy 721 ggcaaaaaaagaatttctgtatcagaccagattttctcagacaaactgttattacctttgt 780
 Db 721 ggcaaaaaaagaatttctgtatcagaccagattttctcagacaaactgttattacctttgt 780
 Qy 781 ggcaggtgattttgtatcagattttagtaataatcatgttacttctgggtggattcaac 840
 Db 781 ggcaggtgattttgtatcagattttagtaataatcatgttacttctgggtggattcaac 840
 Qy 841 accaacaatgaacatgagccagcaagtgatatctgccccacacattctgtcgggaaca 900
 Db 841 accaacaatgaacatgagccagcaagtgatatctgccccacacattctgtcgggaaca 900
 Qy 901 tctgtgcaaaaatttctacactgagccagcgagtgatcttctgtgaactccgggcattt 960
 Db 901 tctgtgcaaaaatttctacactgagccagcgagtgatcttctgtgaactccgggcattt 960
 Qy 961 gactgggggagtgagaccacaaaatctggaaaaatgcaatgcacccaactctctgaagtac 1020
 Db 961 gactgggggagtgagaccacaaaatctggaaaaatgcaatgcacccaactctctgaagtac 1020
 Qy 1021 agactcagagatagcaggtccctcactacagcaatgtggacagaggtcagggactgttca 1080
 Db 1021 agactcagagatagcaggtccctcactacagcaatgtggacagaggtcagggactgttca 1080
 Qy 1081 aatccagaagacgtgaaaaatgctgctctctgaggtgacacacccatctaccataagaat 1140
 Db 1081 aatccagaagacgtgaaaaatgctgctctctgaggtgacacacccatctaccataagaat 1140
 Qy 1141 attctcgaatgggtcactgaggtatttctatctggtgttggatgctctcaccgctatgtac 1200
 Db 1141 attctcgaatgggtcactgaggtatttctatctggtgttggatgctctcaccgctatgtac 1200
 Qy 1201 aatgaataatcactcactgtatgcagcagagagagagacacaccccttcccgagcgggtgag 1260
 Db 1201 aatgaataatcactcactgtatgcagcagagagagagacacaccccttcccgagcgggtgag 1260
 Qy 1261 gccgtattg 1269
 Db 1261 gccgtattg 1269

RESULT 2

AAF45131
 ID AAF45131 standard; cDNA; 2044 BP.

XX AC AAF45131;

XX DT 30-MAR-2001 (first entry)

XX DE Human TANGO 294 cDNA.

XX KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

XX OS Homo sapiens.

XX PN WO200077239-A2.

XX PD 21-DEC-2000.

XX PF 24-MAY-2000; 2000WO-US14858.

XX PR 14-JUN-1999; 99US-0333159.

XX PA (WILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI: 2001-032313/04.
 DR P-PSDB; AAB66065.
 XX
 PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 1; Fig 6; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-P45136 and AAF45138-P45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;
 Query Match 100.0%; Score 1269; DB 22; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 atgttgaacccttgcagacagctgattgtctcacagaaatggaatggctctg 60
 DB 126 atgttgaacccttgcagacagctgattgtctcacagaaatggaatggctctg 185
 QY 61 attctgtggcggtatatttccagagaaatggaattcagtcacatatgccaaactaaagct 120
 DB 186 attctgtggcggtatatttccagagaaatggaattcagtcacatatgccaaactaaagct 245
 QY 121 gtgaccagaaagcattcagaaatattagtgaaatccatccaaactcagctatccctgt 180
 DB 246 gtgaccagaaagcattcagaaatattagtgaaatccatccaaactcagctatccctgt 305
 QY 181 gaggaatatgaatgcgaactgaagatgggtatattctcttcttaacagattccctga 240
 DB 306 gaggaatatgaatgcgaactgaagatgggtatattctcttcttaacagattccctga 365
 QY 241 ggcctagtgcacactaagaagacaggttccagcctgtgtgttactgcagatggccta 300
 DB 366 ggcctagtgcacactaagaagacaggttccagcctgtgtgttactgcagatggccta 425
 QY 301 gttggaggtgctagcaactgatttccaaactgcccacaaatagcctgggcttctctg 360
 DB 426 gttggaggtgctagcaactgatttccaaactgcccacaaatagcctgggcttctctg 485
 QY 361 gcagatgctgttttgacgtgtgattggggaacagcaggggaaacgcctgtgtctcgaaa 420
 DB 486 gcagatgctgttttgacgtgtgattggggaacagcaggggaaacgcctgtgtctcgaaa 545
 QY 421 cacagacactctccatagaccagatgattctgtgggttctcagttatgatgatggct 480
 DB 546 cacagacactctccatagaccagatgattctgtgggttctcagttatgatgatggct 605
 QY 481 aggtttgaccttctcagtgataaactttatttgcagaaaacggccaggaagaaatc 540
 DB 606 aggtttgaccttctcagtgataaactttatttgcagaaaacggccaggaagaaatc 665
 QY 541 tattatgtcgctattcacaggggaccacacatgggctttattgtattttccacatgcc 600
 DB 666 tattatgtcgctattcacaggggaccacacatgggctttattgtattttccacatgcc 725
 QY 601 gactggctcagaaaaacaaatgtatttgccttttagcaccatagccactgttaagcat 660
 DB 726 gactggctcagaaaaacaaatgtatttgccttttagcaccatagccactgttaagcat 785
 QY 661 gcaaaagcccgaggacaaatatttgcctggcagatgatgatgaagggattgttt 720
 DB 786 gcaaaagcccgaggacaaatatttgcctggcagatgatgatgaagggattgttt 845
 QY 721 gcaaaagaaatatttgcctggcagacagatttctcagacaaactgttattacattgt 780
 DB 846 gcaaaagaaatatttgcctggcagacagatttctcagacaaactgttattacattgt 905
 QY 781 ggcaggtgattcttgatcagattgttagtaataatcatgttacttcgggggattcaac 840
 DB 906 ggcaggtgattcttgatcagattgttagtaataatcatgttacttcgggggattcaac 965
 QY 841 accaacaatgaacatgagccagcagcaagtgtatgtctgccacactcttgcctggaa 900
 DB 966 accaacaatgaacatgagccagcagcaagtgtatgtctgccacactcttgcctggaa 1025
 QY 901 tctgtcaaaatattctacacagcagcagcagcagcagcagcagcagcagcagcagc 960
 DB 1026 tctgtcaaaatattctacacagcagcagcagcagcagcagcagcagcagcagcagc 1085
 QY 961 gactgggggagtgagacacaaatctggaataatgcaatcagcacaactcctgttaagtc 1020
 DB 1086 gactgggggagtgagacacaaatctggaataatgcaatcagcacaactcctgttaagtc 1145
 QY 1021 agagtgcagatgacggttccctacagcaatgtggacagagatcagagactggcttca 1080
 DB 1146 agagtgcagatgacggttccctacagcaatgtggacagagatcagagactggcttca 1205
 QY 1081 aatcagaagacagtgaaatgctctctcagtgagtgagcagcacaactcactacataaga 1140
 DB 1206 aatcagaagacagtgaaatgctctctcagtgagtgagcagcacaactcactacataaga 1265
 QY 1141 attcctgaatgggctcagtgatttcatctgtgggtttggatgctcctcaccgtatgac 1200
 DB 1266 attcctgaatgggctcagtgatttcatctgtgggtttggatgctcctcaccgtatgac 1325
 QY 1201 aatgaaatcatcattcattgacgagcagcagcagcagcagcagcagcagcagcagcag 1260
 DB 1326 aatgaaatcatcattcattgacgagcagcagcagcagcagcagcagcagcagcagcag 1385
 QY 1261 gccgtattg 1269
 DB 1386 gccgtattg 1394
 RESULT 3
 AAQ42310
 ID AAQ42310 standard; DNA: 1378 BP.
 XX
 AC AAQ42310;
 XX
 AC AC
 XX
 DT 20-SEP-1993 (first entry)
 XX
 XX RGL precursor.
 DE
 XX Rabbit gastric lipase; RGL; pRGLN2.1; fat; bioconversion;
 KW hydrolysis; transesterification; ds.
 XX
 OS Oryctolagus cuniculus.
 XX
 XX Key Location/Qualifiers
 FH misc_RNA 120..1247
 FT /tag- a
 FT /note- "claim 1; page 10-11"
 XX
 PN EP542629-A.
 XX
 PD 19-MAY-1993.
 XX
 PF 12-NOV-1992; 92EP-0403055.
 XX

PR 13-NOV-1991; 91FR-0013948.
 XX (LJOU) INST RECH JOUVEINAL.
 XX Benicourt C, Blanchard C, Junien J;
 XX WPI; 1993-161080/20.
 DR P-PSDB; AAR37302.
 XX
 XX Rabbit gastric lipase, its precursor and their DNA - useful for
 PT treating conditions linked to gastric lipase deficiency, such as
 PT mucoviscidiosis and pancreatic exocrine insufficiency
 XX Claim 1; Fig 6; 31pp; French.
 XX
 CC mRNA was isolated from rabbit gastric mucosal cells, converted to
 CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
 CC to transform E.coli MW294 and cells screened for reactivity with
 CC probes corresp. to known parts of the RGL sequence. One positive
 CC clone contd. plasmid p70101 which contd. a 1.35 kb PstI sequence.
 CC This fragment was isolated, subjected to PCR and the amplification
 CC prod. ligated with the 2.2 kb BglII-SacI fragment of pRU276 (contg.
 CC a synthetic Tac promoter and transcription stop signal). The
 CC recombinant plasmid was designated pRGLN2.1.
 CC
 SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;

Query Match 38 68; Score 489.2; DB 14; Length 1378;
 Best Local Similarity 65.08; Pred. No. 2.7e-145;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

OY 125 acccagaagcattcatgaattagtagaatacctcaacacacagcgtatccctgtgagg 184
 DB 136 acctgaagtgaattgaattagtcagatgatttctactctgggatacccaagtga 195
 OY 185 atataagtcgcactgaagtcggtatatactcttctgttaacaggattctctgagcc 244
 DB 196 atataagtcgactgaagtcggtatatactcttctgttaacaggattctctgagcc 255
 OY 245 tagtgcacctaagaagcaggttcacgcctgtggtgttactcagcagcgtcctagt 304
 DB 256 agaaaaattcagggaacagagccagagacccgtcgtatttttcagcagcgttgc 315
 OY 305 gagggtgtagcaactggtattcccaactgcgcacacatagcctgggtcttctctgcag 364
 DB 316 catcagctcaaacgtgattctcccaactgcgcacacacagcctgttattctgcag 375
 OY 365 atgctggttttgacgtgtggtggaacagcaggggaaacgcctgtctcgaaaaacaca 424
 DB 376 atgctggttttgacgtgtggtggaacagcaggggaaacacacgtgtccagaagaa 435
 OY 425 agacactctccatagaccaaagtcagttctggccttcagttatgtagatgagcctaggt 484
 DB 436 tatactattccacagactccgtggaattctggccttttagctttgtagaaatggc 495
 OY 485 ttgacctctctgcagtgataaacttttttgcagaaaacggccaggaagatctatt 544
 DB 496 atgacctctccagcccaattgacttcattgttaaaggaaactggacaggaagcgtgc 555
 OY 545 atgctggtatttccagggccacacacatggcctttattgcatatttccaccatgccag 604
 DB 556 atgctggtatttccagggccacacacatggcctttattgcatatttccaccatgccag 615
 OY 605 tggctcagaataaataatgatttttctgttagcaccatagcactgttaagatgcaa 664
 DB 616 tggctcagaataaataatgatttttctgttagcaccatagcactgtggaatgata 675
 OY 665 aaagcccccgggacaaaatttttggctgcgcagatgatcatcaggaggtttgtgca 724
 DB 676 aaagcctgttaacaaactaggtttattctctccaaacatgttcaagattattgtg 735
 OY 725 aaaaagaatttctgatcagaccagatttctcagacaaacttgttattaccttgggcc 784

DB 736 acaaaattattccaccacacaaattctttgatcaatttcttgccactcaagtggttccc 795
 OY 785 agggattcttgatcagattttgtagtaataatcttacttctggttgattcaacaca 844
 DB 796 gtgagacactgaatgaatttgcagcaatgcttatttattctatttctggttgacagc 855
 OY 845 acaatatgaactgagccgagcaagtgatatgtctgcccacactcttctggaacatctg 904
 DB 856 caaactgaacatgagtcgctggtgtagtgtagtgatcaataatccgcgaggaacttcag 915
 OY 905 tgcataatattctacactggagccagcagtgaaattctggtgaactccggcgttgcact 964
 DB 916 ttcaaacatgctgcaactggaccaggctgttaaatctggaattttcaagcttttaatt 975
 OY 965 ggggagtgagacaaataatctgaaaaatgcaaatcagccacactctgtgaaggtacag 1024
 DB 976 ggggaagcccagctcagaatgtagtgcacttcaatcagcccacactccctactaca 1035
 OY 1025 tcagagatatgacggtccctacagcaatgtggacaggggtcagactggtttcaaatc 1084
 DB 1036 tgaacgcacatgaatgtccaattgcagtgtagtggtggaactgactggttgctgacc 1095
 OY 1085 cagaagacgtgaaaatgctgtctctgtggtgagccaaactcactcaccataagaatctc 1144
 DB 1096 ccnaagatgttgacctttgtctccaaactttctaatctcatttaccacaaggagattc 1155
 OY 1145 ctgaatgggtcagctgagattctatctggttgggttggatgctcctcactcagtgataatg 1204
 DB 1156 ttcatataacactggtatttctctggtggaatgaatgctcctcagaagtttcaatg 1215
 OY 1205 aaatcattccatctgatgcagcagaggagga 1234
 DB 1216 aaattattctgatggcgaagataaaa 1245

RESULT 4
 AAN60685
 ID AAN60685 standard; DNA; 1367 BP.
 XX AC AAN60685;
 XX DT 08-JUN-1991 (first entry)
 XX DE Sequence encoding pregastric lipase.
 XX KW Precursor polypeptide; secretion vector; enzyme; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT sig_peptide 47..103 /*tag= a
 FT mat_peptide 104..1243 /*tag= b
 FT
 XX W08603778-A.
 XX PN 03-JUL-1986.
 XX PD 23-DEC-1985; 85WO-GB00599.
 XX PF 21-DEC-1984; 84GB-0032483.
 XX PR 23-DEC-1985; 85WO-GB00599.
 XX PR 01-JAN-1986; 86GB-0019568.
 XX PA (BREW-) BREWING RES FOUND.
 XX PA (TUBB/) TUBB R S.
 XX PI Tubb RS;
 XX WI 1986-182910/28.
 XX DR P-PSDB; RAP60724.

Best Local Similarity 64.2%; Pred. No. 3.5e-139; Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

```
QY 119 ctgtgacccagcagcattcatgaattagtgaaatcccaacatcaagcgtatccct 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ctggaagccctgaagtacatgaacattagtcagatgattacttattgggatacccaa 182

QY 179 gtgaggaatgaagtgcgaatgaagtggtatattctttctgttaacagattccctc 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 atgaagaataatgaagttgtgactgaagatggtattattcttgaagcaatagattccct 242

QY 239 gagcctagtgcaacctgaagaacagagttccaggcgtgtgttactgcagcatggcc 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 atgggaagaataatcagggaacacagggccagagagatgtgtgttttgcagcatgggt 302

QY 299 tagttgaggtgctagcaactgatttccaaacctgcgaacatgctgggcttcattc 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 tgcttgatcagccacaaactgatttccaaacctgcgaacacagccttgcttcattc 362

QY 359 tggcagatgctgtttttgacgtgtgagtggtggaacagcaggggaaacgctgtctcgaa 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 tggcagatgctgtttttgacgtgtgagtggtggaacagcaggggaaacacacctggccgaa 422

QY 419 aacacaagacactctccatagacacagatgattctgggcttcagttatgatgagatgg 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 gaaactgtactattcaccagattcagttgaaattctgggcttcagttatgatgaaatgg 482

QY 479 ctaggttgacctctcgcagtgataaaactttatttgcagaacagggccaggaagaaga 538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 ctataatgaccttccagccacaaactgacttcttgaagaacaaactggacagaagcagc 542

QY 539 tctatttgctgctattcagcagggccaccacctgggctttattgcatctttccaccatgc 598
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 543 tacactatgtggccattccagggccaccacctgggctttattgcttttccaccatc 602

QY 599 cagagctgctcagaaatacaaatgatttttctttagcaccatagcactgttaagc 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 603 ccagctggctaaagaatacaaacctctcctgcttagctgtgtgcccactgtgaagt 662

QY 659 atgcaaaagcccgccagcaaaattttgtgctgcagatgatgatcaaggatgtg 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 663 atacaaaagccctataaacaacttagatttctcctcaatcccttcaagttatat 722

QY 719 ttggcaaaaagaattctctatcagaccagatttctcagacaaactgttatttaccctt 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 ttggtgcaaaaatttctcaccacacaaacttcttctgcatcaattctgtactgaagt 782

QY 779 gtggccaggtgctcttgatcagattttagtaatactgatttactctgttggtgattca 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 783 gctccctgagatgctgaattctcttgcagcaatgcttattataatttggatttg 842

QY 839 acacaaacaatatgaacatgagccgagcaagtgatatgtgccacacactctgtcgtaa 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 acagtaagaactttaacacagctgctgtgattgtatctatcacataatccagcaggaa 902

QY 899 catctgtcaaaaattctacactgagccagcagtgaaattctgttgaaactccggcat 958
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 cttctgtcaaaaactattctcattggaccagcagtggttaagctgtggaaattccaaact 962

QY 959 ttgactggggagtgagacaaaatactggaataatgaataatcagcaactcctgtgaagt 1018
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 atgactgggaagcccgagttcagaatgagtgactatgatcagtcaccaactccctact 1022

QY 1019 acagagtcagagatgacagctcctcagcaatgttgacagaggtcagagctgcttt 1078
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 acaatgtgacagcattgaattgacaaattgcaagtggtggaacggtggaaggaactgtg 1082

QY 1079 caaatccagagagcgtgaaatgctctctctgagtgaccacactcattaccataaga 1138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1083 ctgaccccaagatgtggccttttctcacaacaccccaactcttattaccacaagg 1142

QY 1139 atattctgaaatgggctcagctgatttctatctggggtttggatgctcctcagctatgt 1198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1143 agattcctttttacaatcacttgaccttatctgtggcaatggtgacctccaagaagttt 1202
QY 1199 acaatgaaatcactccactctgat 1220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1203 acaatgacattgtttctatgat 1224
```

RESULT 6

AAT58916
ID AAT58916 standard; DNA; 1367 BP.

XX AAT58916:

AC AAT58916:
XX 19-AUG-1997 (first entry)

XX Human gastric lipase coding sequence.

XX Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;
KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.

XX Homo sapiens.

XX Key Location/Qualifiers
CDS 47..1243
FT /tag= a
FT /product= gastric lipase
FT sig_peptide 47..103
FT /tag= b
FT mat_peptide 104..1240
FT /tag= c

XX W09633277-A2.

PN 24-OCT-1996.

XX 19-APR-1996; 96WO-FR00606.

XX 20-APR-1995; 95FR-0004754.

XX (BIOC-) EIOCEM SA.

PA (LJOU) INST RECH JOUVEINAL.

XX Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
PI Merot B;

XX WPI; 1996-485783/48.

DR P-PSDB; AAW09383.

XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
PT plants - useful for facilitating absorption of fat, as bio-catalysts
PT and for prodn. of bio-fuel.

XX Claim 3; Fig 4; 130pp; French.

XX This is the nucleotide sequence encoding the human pre-duodenal (i.e.
CC gastric) lipase enzyme. The sequence can be used to generate transgenic
CC plants producing recombinant lipase in an enzymatically active form.
CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
CC the mature protein, respectively) can be deleted to form the derivatives
CC designated delta-4 or delta-54 respectively. Plants, or their extracts,
CC expressing the lipases or the truncated derivatives, can be used:
CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
CC in healthy subjects or in patients with inadequate levels of
CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
CC pancreatic insufficiency, in the elderly or in patients undergoing medical
CC treatment which alters fat adsorption; (b) for performing industrial or
CC agricultural reactions, e.g. in processing of fats or in the dairy
CC industry, for hydrolysis or transesterification reactions, etc., where
CC the plant material may provide both enzyme and substrate. The transgenic
CC plants can also be used for biofuel production.

XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
SQ

Query Match 37.08; Score 470; DB 17; Length 1367;
Best Local Similarity 64.2%; Pred. No. 3.5e-139;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtggaccagagcattcatgaattattgtaataatgtaataatcacaacatcaaggtatccct 178
DB 123 ctggaacccctgaagtgactatgaacattatgctagatgattactattggtggaataccaa 182

QY 179 gtgaggaatagagtcgcaactgaagatgggtatattccttctgttaacaggaattcttc 238
DB 183 atgaagaatagagtgactgaagatggttatatttgaagtcaatagaattcctt 242

QY 239 gaggcctagtgaacacctgaagaagaggttccaggccctggtggtgttactcagcatggcc 298
DB 243 atgggaagaaattccgggaatacagagccagagacctgtgtgttttgcagcatggt 302

QY 299 tagtgaggtgctagcaactggaattccaaacctgcccacaataatagctgggttcttc 358
DB 303 tgcctgcatcagccacaaattggaattccaaacctgcgaacaacagccttgcttcttc 362

QY 359 tggcagatgctggttttgagctgtggatggggaacagcagggaacgctggtctcgaa 418
DB 363 tggcagatgctggtttatgagtggtgctgggaacagcagaggaacacacctgggccaaga 422

QY 419 aacaaagacactctccatagaccagagatgagttctggtgttccagttatgatgagatg 478
DB 423 gaaactgtactattccacagattcagttgaattctggtgttccagttatgatgagatg 482

QY 479 ctagggttgccttccctgagtgataaaactttattttgcagaaaacggccaggaaga 538
DB 483 ctgaatgaccttccagccacacatgcacttcattgtaaaagaacactggacagaagcagc 542

QY 539 tctattatgctggctattccagggccaccacattggttattgcaattttccaccatgc 598
DB 543 tacactatgtggccattccaggccacacattggtttattgcttccaccatgc 602

QY 599 caggcctggtcgagaataatcaaatgtattttgcttagcaccatagccactgttaagc 658
DB 603 ccagcctggtcgagaataatcaaatgtattttgcttagcaccatagccactgtgaagt 662

QY 659 atgcaaaaagcccggaacaaattttgtgtgctgcagatatgatatacaagggatgt 718
DB 663 atacaaaagcccttataaacaacttagattgttctcctcaactccttccaaagttatat 722

QY 719 ttggcaaaaagaatttctgtatcagaccagatttctcagacaactgttatttaacctt 778
DB 723 ttggtgacaataattctaccacacaaacttttggataatttctgtactgtaagtgt 782

QY 779 gtggccaggtgattctctgacagatttggtagtaataatcattgttacttctgggtgattca 838
DB 783 gctccctgagatgctgaattctcttgcagcaatgcttatttataatttgggattg 842

QY 839 acaccacaatatgaacataagcagcagagtgatgtatgctgccacaactctgtctggaa 898
DB 843 acagtgaagaacttcaacacagctgctggtgattgtatctatcacataatccagcaggaa 902

QY 899 catctgtcaaaaattctacactggagccaggcagtgaaattctgtgtaactcgggcat 958
DB 903 ctctgttcaaaaactgttccattggaccaggcgtgtaagtctgggaattccaaagctt 962

QY 959 ttgactggggagtgagcaaaaatactggaataatgcaatcagccaactcctgtaaggt 1018
DB 963 atgactggggagtgagcaaaaatactggaataatgcaatcagccaactcctgtaaggt 1022

QY 1019 acagagtcagagatatgacgttccctacagcaaatgtgacagaggtcaggaactggcttt 1078
DB 1023 acaatcgacagcagcatgaaatgtaccaatgtcagtggtggaacggtggcaagcagctgttg 1082

QY 1079 caaatccagagcagctgaaatgtctctctgtgagtgacacacactctatccataaga 1138
DB 1083 ctgaccccccaagatgtgtggccttttgcctccaaaactcccccaactcttatttaccacagg 1142

QY 1139 atattcctgaatgggtcagctgagtgatttcattcctctgggttggatgctcctcaccgtatgt 1198
DB 1143 agattcctttttacaatcacttggaactttattctgtggcaatggatgcccctcagaagttt 1202

QY 1199 acaatgaatacatccatctgat 1220
DB 1203 acaatgacattgtttctatgat 1224

RESULT 7
AAN50385
ID AAN50385 standard; DNA; 1336 BP.
XX
AC AAN50385:
XX
DT 17-JAN-1992 (first entry)
XX
DE Rat prelingual lipase gene.
XX
KW Prelingual lipase; enzyme; EC-3.1.1.3; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
CDS 7..1194
FT /*tag= a
XX
PN GB214233"-A.
XX
PD 16-JAN-1985.
XX
PF 29-JUN-1984; 84GB-0016581.
XX
PR 01-JUL-1983; 83GB-0017989.
PR 05-SEP-1983; 83GB-0023759.
XX
PA (CELL-) CELTECH LTD.
XX
PI Carey NH, Williamson R;
XX
DR WPI; 1985-014450/03.
XX
P-PSDB; AAP50322.
XX
PT New lingual lipase protein for treatment of lipase deficiency - also
PT new pre-lingual lipase protein and related products
XX
PS Disclosure; Fig 5; 15pp; English.
XX
CC This sequence may be expressed in a transformant host organism using
CC a vector plasmid in order to produce a prelingual lipase protein.
CC The expressed protein may be used for the treatment of lipase
CC deficiency.
XX
SQ Sequence 1336 BP; 409 A; 285 C; 263 G; 379 T; 0 other;

Query Match 36.6%; Score 464.4; DB 6; Length 1336;
Best Local Similarity 63.4%; Pred. No. 2.1e-137;
Matches 711; Conservative 0; Mismatches 411; Indels 0; Gaps 0;

QY 119 ctgtggaccagagcattcatgaattattgtaataatgtaataatcacaacatcaaggtatccct 178
DB 80 ctggaacccctgaagcaaatgaatatttagtcagatgataacttacttgggttatccat 139

QY 179 gtgaggaatagagtcgcaactgaagatgggtatattccttctgttaacaggaattcttc 238
DB 140 gtcaagaatatgaattgttactgaagatggtctacattctgggggtctacagaattcttc 199

QY 239 gaggcctagtgaacacctgaagaagcaggttccaggcctggtgttactcagcatggcc 298
DB 200 atgggaagaataattctgaaataataggaagacctgtgggtgatttgcagcatggtt 259

Qy 425 agaacctctccatagacacagatgagttctgtggttttcagttatgatgagatgaggt 484
 Db 326 tgtactactccgcagactccgtcgaaattctgtggttttcagttatgatgagatgaggt 385
 Qy 485 ttgacctctctcagtgatgaacttttttttcagaaacggccaggaagaagatctatt 544
 Db 386 atgacctctccgcacacactctcagttcctcttgaaagaaacggccaggaagaagatctatt 445
 Qy 545 atgtcgggtattctacagggccaccacattgttattgtttttccaccatgccagagc 604
 Db 446 acgttgccattccaggccaccacattgttattgtttttccaccatgccagagc 505
 Qy 605 tggctcagaaatacaaaatgtatttctgttttagcaccacatagccactgtttaagatgcaa 664
 Db 506 tggcgaacaggtacaaacactctctcagttcctcttgaaagaaacggccaggaagaagatctatt 565
 Qy 665 aaagcccgccgacacaaattttgtgtcgcagatatgatcatgaaggatgtttggca 724
 Db 566 aaacctgtttaacaaactctcgtcgtcccttctgttctccttcaagcttattttggaa 625
 Qy 725 aaaaagaatttctgtatcagaccagattttctcagacaaactgtttatttaccctttgtggcc 784
 Db 626 acaaaattctaccacacacactttctgatcaatttctgcacacaggtatgctccc 685
 Qy 785 aggtgattcttgatcagatttctgtaataatcatttcttctctgttgggttaccacacca 844
 Db 686 gcgagacggttgatctctctgcagcaacgcccgtttatctatttggattgacacta 745
 Qy 845 acaatgaatcagacgagcgaagtgtatgtctgcccacactcttctgtggaacatctg 904
 Db 746 tgaacttgaaatcagatgctgtggtatgtctgtctcacaatactcagcaggaacatctg 805
 Qy 905 tgcgaataattctacactggcagcaggtgaattctgtggaactccggcatttgact 964
 Db 806 ttcaagacgtgtctccactggtccagcgtgttaagtctgggaagttccaaagctttgact 865
 Qy 965 ggggaggtgagacacaaataatggaaaaatgcataatcagccaaactctgtgaaggatcagag 1024
 Db 866 ggggagccaggttcagacatgatgcactatcagagatgctcctcactacacacc 925
 Qy 1025 tcagagatatcaggtccctcagacatgtggacagaggtcagactggttccaaatc 1084
 Db 926 tgcagacatgcatgctcgaatcagctgctggaacggtggcaagactgtgtgcccagc 985
 Qy 1085 cagaagacgtgaaatgctctctgtgagtgacacaaactctatctacataagaatattc 1144
 Db 986 ctacagatgtgacctttgtcttccaaagctcccaatctcatttaccacaggaagattc 1045
 Qy 1145 ctgaatgggtcagctggtattctatctgtggttttgatgctcctcactcagcgtatgtacaatg 1204
 Db 1046 ctccctacaatcacttgactttatctgtggccatggtgcccctcagcgttttacaatg 1105
 Qy 1205 aaatcatccatctgatg 1221
 Db 1106 aaattgtttccatgatg 1122

RESULT 9
 AAT58915
 ID AAT58915 standard; cDNA to mRNA; 1528 BP.
 XX
 AC AAT58915;
 XX
 DT 19-AUG-1997 (first entry)
 XX
 DE Dog gastric lipase coding sequence.
 XX
 KW Dog; duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; ss.
 OS Canis familiaris.

FH Key Location/Qualifiers
 FT CDS 1..1140
 FT /*tag- a
 FT /product= gastric lipase
 XX
 PN W09633277-A2.
 XX
 PD 24-OCT-1996.
 XX
 PF 19-APR-1996; 96WO-FR00606.
 XX
 PR 20-APR-1995; 95FR-0004754.
 XX
 PA (BIOC-) BIOCEM SA.
 PA (LJOU) INST RECH JOUVEINAL.
 XX
 PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 PI Merot B;
 XX
 DR WPI; 1996-485783/48.
 DR P-PSDB; AAW09382.
 XX
 PT Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 XX
 PS Claim 2; Fig 1; 130pp; French.
 XX
 CC This is the nucleotide sequence encoding the dog pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 12 bases of this sequence (encoding amino acids
 CC 1-4 of the protein encoded by this sequence) or bases 1-162 of this
 CC sequence (encoding amino acids 1-54) can be deleted to form the
 CC derivatives designated delta-4 or delta-54 respectively. Plants, or
 CC their extracts, expressing the lipases or the truncated derivatives, can
 CC be used: (a) as pharmaceuticals or food to facilitate absorption of fat,
 CC either in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX
 SQ Sequence 1528 BP; 392 A; 387 C; 329 G; 420 T; 0 other;

Query Match 35.0%; Score 444.2; DB 17; Length 1528;
 Best Local Similarity 62.8%; Pred. No. 6.1e-131;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 Qy 125 acccagaagcattcatgaattagtgaaatcattcaacatcaaggctatccctgtgag 184
 Db 26 accctgaagtgaacctgaataatgaatcagatgatacctactggggataccacagctgag 85
 Qy 185 aatgtgaagtgcgaactgaagtgggtatattcttctgttaacaggattctctcaggcc 244
 Db 86 aatgtgaagtgtgaccgaagacggttatattctctgggacgacagaattctcttgga 145
 Qy 245 tagtgcaaccttaagaagacaggttccagcctgtgtgttactcagacatggcctagtgt 304
 Db 146 ggauaaattcagagaataatagccggagacctgtgtgcatttttgcaacacggtttgctg 205
 Qy 305 gaggtgtcagcaactggatttccaaacctgcccacaaataagctgggtcattcttgagcc 364
 Db 206 catcagccacaaactggatctctccaaacctgcccacaaagcctggccttcatctctgccc 265
 Qy 365 atgtcgggttttgacgtgtggatggggaaacagcagcagggaaacccctgtgtctcgaacaca 424
 Db 266 acgcccgggttacgacgtgtggtggtgggaacagcaggggggaacacctgggcccaggaggaatc 325

QY	425	agacactctccatagaccagaatgagttctgtggcttccagtattgatgagatggtcaggt	484
Db	326	tgtactactcgcgcacccctcgcaattctgtggtcttccagcttgcagagatggcttaat	385
QY	485	ttgaccttctcagtgatataaacttatttgcagaaacaggccaggaaagatctatt	544
Db	386	atgaccttccgcgccaccttgacttcatcttgaaagaaacaggacaggacaagtacact	445
QY	545	atgtcggctattccagggcaccaccattgggctttattgcattttccaccatgcacagc	604
Db	446	acgttggccattccagggcaccaccattggttctcatcgcttttccaccaatcccaagc	505
QY	605	tggctcagaaatacaaatgtatttggcttttagcaccatagccaactgttaagcatgcaa	664
Db	506	ttgcgaaacggtacaaacctctatgatgatgctcccgcttgcaccgtgaagtacacg	565
QY	665	aaagcccccgggaccaaaatttggctgcgcagatatgatcaagggtattgttgcca	724
Db	566	aaaccctgttaacaaactcatgtcgtccctctgttctcttcaaagcttatatttgaa	625
QY	725	aaaaagaattctgtatcagaccagatttctcagacaacttgttattaccctttgtggcc	784
Db	626	acaaaataattaccaccacacctctctgtatcacaattctcgccacggagtatgctcc	685
QY	785	agtgattcttgatcagattttagtaataatactgttacttctctgggtgggattcaacacca	844
Db	686	ggagacggtgatctctctcgcagcaacgcctgttatcaattgttgatttgacacta	745
QY	845	acaatatgaacatgagccgagcaagtgatatgtctgcgccacactcttctgtggaacatcg	904
Db	746	tgaacttgaaactgagctgtggatgtgtatctctgtcaataaattccagcagggaacac	805
QY	905	tgcataaattctcacactggagccaggccagtgaaattctgttgaaactcgcgcgcatcttgact	964
Db	806	tcagaaactgtctcactgtgtccaggctgttaagtctgggaagtccaagcttttgact	865
QY	965	ggggggagtgagaccaaaactcggaaaactgcatacagccaactcctgttaaggatcacag	1024
Db	866	gggggaagccaggttcagaaactgtatcactatcatcagagcatgctcctctactacaacc	925
QY	1025	tcagagatatgacggtccctacagcaatgttgagcaggagggtcaggactggcttccaatc	1084
Db	926	tgcagacacatgcatgtgccaatcgcagtgtaggaacgggtggcaacgctgtgtgcccagcc	985
QY	1085	cagaacagctgaaatgctctctctgaggttgacaaacctcatctaccataagaataatc	1144
Db	986	ctcaogatgttgacctttgtctccaagctccccaactctcatctaccagggaagattc	1045
QY	1145	ctgaaatgggctcacgtggatttcactctgggtttggatgctctccacgatgtacaatg	1204
Db	1046	cctcttaacaataccttggaacttactgtggccatggatgccccctcaagcgggttcacatg	1105
QY	1205	aaatcatccatctgatg	1221
Db	1106	aaattgtttccatgatg	1122

RESULT 10

AAQ68389

ID AAQ68389 standard; DNA; 1531 BP.

XX

AC AAQ68389;

XX
DT 20-FEB-1995 / 14:04:00

DT 20-FEB-1995 (first entry)
XX

Canine gastric lipase (Met-CGL)

XX

XXXXXXXXXXXX XXXX (net con):

KW Canine gastric lipase; CGL; dog; fat; mucoviscidosis;

KW bioconversion; exocrine pancreatic insufficiency; ss.

XX

OS Canis familiaris.
yy

XX	Key	Location/Qualifiers
FH		

FT	1..1143.	
FT	/ *tag= a	
PN	WO9413816-A.	
XX		
PD	23-JUN-1994.	
XX		
PF	16-DEC-1993; 93WO-FR01260.	
XX		
PR	16-DEC-1992; 92FR-0015201.	
XX		
PA	(LJOU) INST RECH JOUVEINAL.	
XX		
PI	Benicourt C, Blanchard C, Junien J;	
XX		
DR	WPI; 1994-217890/26.	
DR	P-PSDB; AAR56871.	
XX		
PT	Recombinant canine gastric lipase and nucleic acid encoding it -	
PT	are used for improving absorption of ingested fat, treating	
PT	mucoviscidosis etc. and in enzymatic bio-conversions	
XX		
PS	Claim 2; Fig 8; 52pp; French.	
XX		
CC	The sequence given below is the sequence of figure 8, altered	
CC	according to the amendments described on page 2 of the appended	
CC	letter.	
CC	CCGL is used to improve absorption of ingested fat, in healthy and	
CC	sick patients (e.g. having altered levels of gastric lipase); to	
CC	treat conditions associated with insufficiency (or lack) of lipases,	
CC	esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.	
CC	where immobilised, for bioconversions, e.g. hydrolysis or	
CC	transesterification (other mammalian gastric lipases, or derivs.,	
CC	can be used in this application).	
XX		
SQ	Sequence 1531 BP; 395 A; 386 C; 329 G; 421 T; 0 other:	

	Query Match	35.0%	Score 444.2;	DB 15;	Length 1531;
	Best Local Similarity	62.8%;	Pred. No; 6.1e-131;		
	Matches' 689;	Conservative 0;	Mismatches; 408;	Indels 0;	Gaps
QY	125	accagaagcattcatgaatatattagtgaatcatccaacataccaaggctatccctgtgagg	184		
Db	29	acctgaagtgaacctgaatataaagtcagatgatcaccttactgggatacccagctgagg	88		
QY	185	aatttgaaatcgcaactgaagtgggttatatccctttctgttaacaggattctctgaggcc	244		
Db	89	aatttgaaagtctgcgcgaagcggttatctctggatcgacagaattctcttagggga	148		
QY	245	tagtgcacctgaagaacaggttccaggccctgtgggtttactgaagcatgcccctagtg	304		
Db	149	ggaataattcagaataataggccggagacctgttgtcatttttgcaaacaggtttgtctcg	208		
QY	305	gaggtgctagcaactggatttccaacctgcccaacaatagctgtgggtcttactctggcag	364		
Db	209	catacgccaaaactggattctccaacctgcccaacaacagcgtggccttcactctggccg	268		
QY	365	atgctggttttgacgtgtggatggggaacagcaggggaaacgccttggtctcgaacaaca	424		
Db	269	acgccgggttacgcgtgtgtggctggggaaacagcaggggcaacacctgggcccaggaggaatc	328		
QY	425	agacactctcatagaccaagatgagttctgggcttctgatttatatagatggtcaggt	484		
Db	329	tgtactactgcgcgcactcgtcgaaattctggcttcttcagcgtttgacgagatggtcuaat	388		
QY	485	ttgaccttctctgcagtgaataactttattttgcagaaaaacgccccgaaaaagatctatt	544		
Db	389	atgaccttccccgccaccattgacttctcatctttaagaagaaaacgggacagacagctacact	448		
QY	545	atgtcggtctattcaaggcgacacacatgggttttattgtcatttttccacatgcccagagc	604		
Db	449	acgttggcatttcccaggcaccacatttggtttcatcgcctttttccacaatcccagac	508		

CC disease), tumours, bacterial, fungal or viral infections; depression and
 CC psoriasis. AA45926 to AA45931 represent linker variants which are given
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 572 BP; 179 A; 109 C; 124 G; 156 T; 4 other;

Query Match 15.2%; Score 193.2; DB 21; Length 572;
 Best Local Similarity 54.1%; Pred. No. 2.8e-51;
 Matches 288; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

Oy 123 ggaccaccagcattcatgaattagtgaaatcatccaacatcaaggctatccctgtga 182
 Db 123 gaatccctgaagctcacatgaattagtgagattattaaacactgggtattccaagcga 182

Oy 183 ggaatatgaagtcgaactgaagatgggtatattcttctgttaacaggattctctcag 242
 Db 183 anaatatgaggttgactgatgtgttacattcttccaatttaacccaattctctcag 242

Oy 243 cctagtccaacctgaagacaggtttccaggcctgtgtgtttactgcagcatggcctagt 302
 Db 243 gaagaacaatgctaattagttcagccccaagatggttagtattttgtcagcatggctgt 302

Oy 303 tggagggtgctagcaactgatttccaacctgccaacaataagctgggtcttcttctgc 362
 Db 303 tgcacacctggagcagggtgttccaatccgctgtccaacagcctggcctctcatctagc 362

Oy 363 agatgctggtttgacgtggtgaggggaacagcaggggaacgcctggtctcgaaaca 422
 Db 363 agatgctggttatgtgtggtgaggggaacagcaggggaacgcctggtctcgaaaca 422

Oy 423 caagacactctccatagacaagatgagttctggtggtttcagtttatgatgagtgctag 482
 Db 423 cgtggcctcaaccnattctaaagaattctgggatttttagtttgatcaaatgataaa 482

Oy 483 gttgacctctctcagtcagtaaaactttattttgcagaaacagggccaggaagaaatcta 542
 Db 483 atagacctctcagtcacatttatttctgtgataaaacagggacaaagcagattta 542

Oy 543 ttatgctggctatttcacagggccaccaca 571
 Db 543 ctacattggccattctcaaggaaactcga 571

RESULT 15

AA234958
 ID AA234958 standard; cDNA; 1718 BP.

XX
 AC AA234958;

XX 28-FEB-2000 (first entry)

XX Soybean acid triacylglycerol lipase cDNA.

XX Triacylglycerol lipase; soybean; fatty acid; seed oil;
 KW vegetable oil; transgenic plant; ss.

XX Glycine max.

XX Key Location/Qualifiers
 FT 184..1416
 CD5 /*tag= a

XX W09955883-A2.

XX 04-NOV-1999.

XX 29-APR-1999; 99WO-US09280.

XX 30-APR-1998; 98US-0083688.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PI

XX Cahoon E3, Cahoon RE, Kinney AJ, Rafalski JA;

DR WPI; 2003-062036/05.

XX P-PSDB; AAY32309.

XX Novel plant triacylglycerol lipase polynucleotides used to alter the
 PT level of the enzyme in transgenic plants

XX Claim 2; Page 48; 65pp; English.

XX This is the nucleotide sequence of a contig assembled from the
 CC entire cDNA insert in clone ssl.pk0022.al and a portion of the
 CC cDNA insert in clone sdp3c.pk004.n3 encoding the entire soybean
 CC acid triacylglycerol lipase (TAGL) (see AAY32309). The clones were
 CC isolated from soybean developing pod and seedling (5-10 day post
 CC germination) cDNA libraries. Novel acid and neutral TAGL polypeptides
 CC (see AAY32301-17) and polynucleotides (see AAY34950-66) from corn,
 CC Catalpa, rice, soybean and wheat tissues are disclosed. The enzymes
 CC may be prepared recombinantly and used to raise antibodies, which
 CC are used for detecting the enzymes in situ in cells or in vitro in
 CC cell extracts. The polynucleotides may be used to create transgenic
 CC plants in which the TAGL levels are present at higher or lower levels
 CC than normal, or in cell types or developmental processes where they are
 CC not normally found. This would alter the level of triacylglycerol and
 CC cholesterol esters found in those cells. Accumulation of fatty acids
 CC with unusual structures may be a positive phenotype in plants used for
 CC foods. In addition, it may be desirable to eliminate expression of TAGL
 CC genes for certain applications. TAGL enzymes may also be useful for the
 CC processing of plant seed oils and for the development of novel seed
 CC oils. The TAGL enzymes can also be used as targets to facilitate the
 CC design and/or identification of inhibitors of those enzymes that may be
 CC useful as herbicides. This is desirable because inhibition of the
 CC activity of either of the enzymes could lead to an inhibition of plant
 CC growth. The polynucleotides also serve as a source of probes and
 CC primers, which are useful for genetic mapping, as markers for traits
 CC linked to those genes, and to isolate homologous sequences from other
 CC species.

XX Sequence 1718 BP; 492 A; 353 C; 344 G; 529 T; 0 other;

Query Match 7.9%; Score 100; DB 21; Length 1718;

Best Local Similarity 55.4%; Pred. No. 2.5e-21;

Matches 237; Conservative 0; Mismatches 185; Indels 6; Gaps 2;

Oy 166 caaggctatccctgtgagggaatgaagtcgcaactgaagatgggtatctcttctgtt 225
 Db 337 catggatacaagtgctcaagaacacgaggttacacatgtaggttacattctgagcctg 396

Oy 226 aacagattcctcgaggcctagtgcacctaagaagacaggttccag---gcctgtggtg 282
 Db 397 caaaggatcccaagaagtcgaggttaaaagcagtgggagtggaagaaagcaaccagtg 456

Oy 283 ttactgcagcatggcctagttgaggtgctagcaactggatttccaccctgcccaaat 342
 Db 457 gttatacaacatgagttctcttagatggtgatgacatgggtcttctaaacccaccagagcaa 516

Oy 343 agcctgggtcttcttcctgagcatgctggtttgacgtggtgaggggaaacagcagggga 402
 Db 517 gatctgccgtgtgatttagctgataatgattgacgtggtgattgcaaacacagaagga 576

Oy 403 aacgctgtgctcgaaaacacagacactctccatagacccaagatggttctggtcttc 462
 Db 577 accagatagtcgcccagacacatctcatgtggacccctctagcagcctattggaattgg 636

Oy 463 agttatgatgagatggcctaggtttgacctctcctgcagtgataaacttattttgcagaaa 522
 Db 637 tcttggtgagaaacttctctctctatgatttccctcggtgttttaattatgtgttcagcaa 696

Oy 523 acgggcaggaagaaatctattatgctgcgtattctacacagggccaccaccatgggtttatt 582
 Db 697 acggggcag---aagatcaattacgttggccattcttgggaactttggtagctttggca 753

Oy 583 gcatttc 590
| ||||
Db 754 tccttc 761

Search completed: August 24, 2001, 22:42:12
Job time: 7433 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 21:22:40 : Search time 2518.61 Seconds
(without alignments)
4762.814 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgcaag.....gacgggtgagccggtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:
12: gb_est12:
13: gb_est13:
14: gb_est14:
15: gb_est15:
16: gb_est16:
17: gb_est17:
18: gb_est18:
19: gb_est19:
20: gb_est20:
21: gb_est21:
22: gb_est22:
23: gb_est23:
24: gb_est24:
25: gb_est25:
26: gb_est26:
27: gb_est27:
28: gb_est28:
29: gb_est29:
30: gb_est30:
31: gb_est31:
32: gb_est32:
33: gb_est33:
34: gb_est34:
35: gb_est35:
36: gb_est36:
37: gb_est37:
38: gb_est38:
39: gb_est39:
40: gb_est40:
41: gb_est41:
42: gb_est42:
43: gb_est43:
44: gb_est44:
45: gb_est45:
46: gb_est46:
47: gb_est47:

44: em_esthum10:
45: em_esthum11:
46: em_esthum12:
47: em_esthum13:
48: em_esthum14:
49: em_esthum15:
50: em_esthum16:
51: em_esthum17:
52: em_esthum18:
53: em_esthum19:
54: em_esthum20:
55: em_esthum21:
56: em_esthum22:
57: em_esthum23:
58: em_esthum24:
59: em_esthum25:
60: em_esthum26:
61: em_esthum27:
62: em_esthum28:
63: em_estin1:
64: em_estin2:
65: em_estin3:
66: em_estin4:
67: em_estin5:
68: em_estomi:
69: em_estom2:
70: em_estov1:
71: em_estov2:
72: em_estpl1:
73: em_estpl2:
74: em_estpl3:
75: em_estpl4:
76: em_estpl5:
77: em_estpl6:
78: em_estpl7:
79: em_estpl8:
80: em_estpl9:
81: em_estpl10:
82: em_estro1:
83: em_estro2:
84: em_estro3:
85: em_estro4:
86: em_estro5:
87: em_estro6:
88: em_estro7:
89: em_estro8:
90: em_estro9:
91: em_estro10:
92: em_estro11:
93: em_estro12:
94: em_estro13:
95: em_estro14:
96: em_estro15:
97: em_estro16:
98: em_estro17:
99: em_estro18:
100: em_estro19:
101: em_estro20:
102: gb_est25:
103: gb_est26:
104: gb_est27:
105: gb_est28:
106: gb_est29:
107: gb_est30:
108: gb_est31:
109: gb_est32:
110: gb_est41:
111: gb_est42:
112: gb_est43:
113: gb_est44:
114: gb_est45:
115: gb_est46:
116: gb_est47:


```

Db 241 TTTTATACAGACTAGATTTTTCAGGAGCTTTTATTTATTTACC-TTGTGGCCAGATGATT 299
Qy 793 cttgatcagattttagtaatatcatgttacttctgtggtgattcaacacccaacaatatg 852
Db 300 CTCGACCAAAATCTGCAGCAACATCATCTTACTCTCTGGGAGGTTTAAACACAAACAATATG 359
Qy 853 aacatgagccagagaagtgtatctgctgcccacactcttctggaacatctgtgcaaaat 912
Db 360 AACATGAGCAGAGAAATGTGTGTGCCCCATACACTGTCTGGAAGCTGTGCGAAGAC 419
Qy 913 attcactagagccagggcagtgaaattctgtgaaactccgggcatttgcactggggagt 972
Db 420 ATTCTCACTCGAGCCAGCCGCGTGAATCTGGGAACTTCTGCTTTGACTGGGGAAGT 479
Qy 973 gagacccaaaaatctggaaaaatgcaatcagccaaactcctgtgaaggtcacagagtgcagat 1032
Db 480 GAGACCAAAAATCAGGAGAAATGTAATCAGCCAACTCCTATAAGGTACAAAGTTCGAGAT 539
Qy 1033 atgacggtccctacagcaatgtgacagagggtcaggactggttccaaatccagaagac 1092
Db 540 ATGATGGTCCCAACAGCAATGTGACTGGAGGTCAAGACTGG-TTCAAAATCCAGATGAT 598
Qy 1093 gtgaaaatgctctctctgaggtgaccaaactcctcattaccataagaatattccctgaatgg 1152
Db 599 GTGAACAATTTACTTCTGAAGTACCAAACTCATCTACCAACAGACATCTCTGAGGTG 658
Qy 1153 gctcagctggatttcattcctctggttggatgctctccacgg 1193
Db 659 GGTATGTGGA-TTCATCTGGGGGGTGGATGCCCTCAGCG 698

RESULT 3
BF141567 1079 bp mRNA EST 24-OCT-2000
LOCUS 601788654F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5'
DEFINITION mRNA sequence.
ACCESSION BF141567
VERSION BF141567.1 GI:10980607
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1079)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
National Library Arrayed by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9263 row: 1 column: 14
High quality sequence stop: 663.
Location/Qualifiers
1..1079
/organism="Mus musculus"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4016413"
/clone.lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPT6; Site:1: NotI;
Site:2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NTH"
312 a 267 c 263 g 237 t
BASE COUNT

```

ORIGIN

Query Match 39.28; Score 497.6; DB 145; Length 1079;
 Best Local Similarity 84.18; Pred. No. 3.4e-140;
 Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;

```

Qy 256 aagsagacaggttccaggcctgtggttactgcagcatggcctagttggagggtgctagc 315
Db 20 AAGAAGGAAGGATCCAGGCCAGTGGTGTACTGCAACATGCTCTCTTTGGGGATGCTAGC 79
Qy 316 aactggatttccaaacctgcccacaatacagcctgggtctcattcttggcagatgctggttt 375
Db 80 AACTTGGATTTTCCAACTGCCCAACAACAGCCTGGGTTTTTATTTCTGGCAGATGCAGTTTT 139
Qy 376 gacgtgtggtggaagaaacagcaggggaaacgcc-tggctctgaaacacacagacactctc 434
Db 140 GATCTGTGGATGGGAACAGCAGAGGAANACACCCTGGTGTCTCGGAAGCACAAGACCTCTC 199
Qy 435 catagacccaagatgagttctgggttttcagttatgatgagatggctaggtttgaccttcc 494
Db 200 CATAGATCAAGATCAGTTCGGGCTTTCAGTTATGATGAATGGCTAGGTTGACCTTCC 259
Qy 495 tgcagtgataaaacttattttgcagaaacccggccagagaaagatctctattatgtcgcta 554
Db 260 AGCTGTGATAAACTTTATCTTACAGAAACGGCCAGAAAAAGGTCTATTATGCGGCTA 319
Qy 555 ttcacagggcaccacacattgggtcttattgcatttccaccatgccagagctggctcagaa 614
Db 320 CTCACAGGGCACCACCATGGGCTTTATTTGCAATTTTCCACAATGCCAGAGCTAGCTCATAA 379
Qy 615 aatcaaaatgtattttcttttagcaccatagccactgtttaagcatgcaaaaagcccg 674
Db 380 AATCAAAATGTATTTGGCTTTAGCTCTATAGCCACTGTTAAATATGCAAGAGTCTCTGG 439
Qy 675 gaccaatttttctgtgtccagatatgatcaagggtatgtttggcagaaagaaatt 734
Db 440 TACCAAAATTTCTGTCTGCCAGATATGATGATCAAGGTATTATTTTGCGACACAGAGTT 499
Qy 735 tctgtatcacacagatttctcagacaaactgtttatttacccttggccagggtattct 794
Db 500 TTTTATCAGACTAGATTTTTCAGGCAGCTTTTATTATTTACCTTTTGGGCGAGATGATTCT 559
Qy 795 tgatcagattttagtaatatcatgttacttctgtgtggtgattcaacacccaacaatatgaa 854
Db 560 CGACCAATCTGCAGCAACATCATCTTACTCTGGGAGGTTTAAACACAAACCATATGAA 619
Qy 855 catgagccagca--agtgtatatgtgcccacactcttctgtggaacatctgtg---caa 909
Db 620 CATGAGCAGACCAACATGTGCTCTGTGCCCAATAAACCTGTGGAACGCTCTTGTGCGCA 679
Qy 910 aattattcactgagccagg-cagtgaaattctggtg-aactccgggcatttgcactggg 967
Db 680 ACCATTCTCACTCGAGCCAGCCGGTGAATTTGGGGCACCTTGTGCTGCTTGTACTGTG 739
Qy 968 ggaagtgaaccacaaatc 985
Db 740 GGCAAGCGACAAACCAAGC 757

```

RESULT 4

```

BF143546
LOCUS BF143546 669 bp mRNA EST 24-OCT-2000
DEFINITION 601790896F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021870 5',
mRNA sequence.
ACCESSION BF143546
VERSION BF143546.1 GI:10982586
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 669)
REFERENCE

```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs@email.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 cDNA distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM9277 row: o column: 23
 High quality sequence stop: 659.
 Location/Qualifiers
 1. 669
 /organism="Mus musculus"
 /strain="C28H II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4021870"
 /clone_lib="NCI_CGAP_Lu30"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; transgenic model WNT-1, expression driven by
 MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
 dt. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 140 c 165 g 169 t
 ORIGIN
 Query Match 37.2% Score 471.8; DB 145; Length 669;
 Best Local Similarity 85.9% Pred. No. 2e-132;
 Matches 535; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
 QY 1 atgttggaacctgtgcaagacagtgattgtctccacacagaatggaaatgtggtctctg 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 48 ATGTCAGAAATCTTGTCAAGAGTGTGGACTGTTTCGCACAGAGTGGAGATATGGCTCTG 107
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 61 attctgtggtcgtatgtctccagagaatgtgaatcagatcacatagcgaactaaagct 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 108 ATTCGTGTAGCATATTACTCCAAAGAAATGTGACTCGGCACATTTGCCACAGAAAGCT 167
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 121 gtggaccagaagcattcatgaatattagtgaaatcatccaaatcagaggtatccctgt 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 168 GCGGATCCAGAAAGCATTCATGATGTTAGCGAAATCATCAACACAGAGGTTATCCAGT 227
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 181 gaggaatatgaatcgcaactgaagtgggtatattctctctgttaacaggattctctga 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 228 GAGGAGTATGAATGTGCAACAGAGATGGGTATCATCTCTGTGTGAACAGAAATCCCTCGG 287
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 241 ggcttagtgaacctgaagacaggtgtccaggcctgtgtgttactgagcagatggccta 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 288 GCACAGACAGGTTAAAGAGGAGGATCCAGGCCAGTGTGTGTACTGCAATATGGTCTT 347
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 301 gttggaggtgtagcaactggatttccaaacctgcccacaaatagcctgggcttcaattctg 360
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 348 CTGTTGGGATGCTAGCAACTGGATTTCACACTGCCCCAACACAGCCTGGGTTTATTCTG 407
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 361 gcagatgctggttttgcgtgtgtgagtgagggaacagcaggggaacacgctggtctgaaaa 420
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 408 GCAGATCAGG-ATTGATGTGTGGATGGGAACACAGCAGAGAAACACCTGGTCTCGGAAG 466
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 421 cacagacacttctccatagaccagatgattcttggtcttctcagttatgatgatgctt 480
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 467 CACAGACCCCTCTCCATAGATCAGATGAGTTCTGGGCTTTCAGTTAATGAAATGGCT 526
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 481 aggttgaccttctcgtcagtgataaaactttattttgcagaaaaacggccaggaagaagatc 540
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 527 AGGTTTGACCTTCCAGCTGTGATAAACTTTATCTTACAGAAAGCGGCCAGAAAGGTC 586
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 541 tattatgtcggctattcacaggcgccaccacattgggtttatttgcattttccacattccca 600

Db 587 TATTATGTGGCTACTACAGGCGCACCATGGCTTTATTCATTTCACAATGCCA 646
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 601 gagctggctcagaataatcaaaat 623
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 647 GAGCTAGCTCATACATCCAAAT 669
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 5
 AKO10093 LOCUS
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 library, clone:2310067K20, full insert sequence.
 AKO10093
 VERSION
 ACCESSION
 KEYWORDS CAP trapper.
 SOURCE
 ORGANISM
 Mus musculus
 clone:2310067K20.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 Carlinici, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Methods Enzymol. 303, 19-44 (1999)
 2 (sites)
 Carlinici, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 3 (sites)
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T.,
 Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,
 Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsuura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 4 (sites)
 The RIKEN Genome Exploration Research Group Phase II Team and
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 (bases 1 to 1350)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
 Handachi, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,
 Imatani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
 Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
 Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
 Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
 Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sojabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
<http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome

**JOURNAL
MEDLINE
REFERENCE
AUTHORS**

Genome Res. 10 (10), 1617-1630 (2000)
20499374
3 (sites)

3 (sites)
Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Akizawa, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T.,
Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T.,
Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsuura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system—384-format
sequencing pipeline with 384 multicapillary sequencer

TITLE

**JOURNAL
MEDLINE
REFERENCE
AUTHORS**

sequencing pipeline with 384 multiplexed samples
 GenBank Res. 10 (11), 1757-1771 (2000)
 20530913
 4 (sites)
 The RIKEN Genome Exploration Research Group Phase II Team and
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 (bases 1 to 1347)

5 (bases 1 to 1347)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Carinci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
 Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,
 Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
 Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
 Miyazaki, A., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
 Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
 Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, I., Toyota, T., Yamamura, T., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

TITLE
JOURNAL

Submitted (10-Jul-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT

FEATURES

```

Location/Qualifiers
1..1347
/organism="Mus musculus"
/strain="C37BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1895801"
/db_xref="MGD:MGI:1914967"
/clone="2310068005"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
51..1338
/note="putative"
/codon_start=1
/protein_id="BAB26711.1"
/db_xref="gi:12845335"
/translation="MWLLLVTSVLSAFGGAHGLFGKLGPKNPANNMVNSQMITTYWYGV

```



```

/clone="2310069p19"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/notes="putative"
/codon_start=1
/db_xref="GI:12845372"
/translation="MWLLLVTSVLSAFGAGHGLFGKLGKPNKNNVNSOMITWYCP
SEYEVVTEDEGVLGVRIPIYCKKNSENIGKRPVAYLOHGLIASATNITNPNLSLA
FILADAGVWLGNSGRNKNYISPDSEVFNAFSDMAKYDLPTAIDFIVQKT
GEKIHVGHSDGTIGFTAFSTPALAKKIRFVALADPIATVTKTESPFKLSLIPK
FLKVFNGKMPHNYLDQFLGTEVCSRELLDLCSNALFICFGDFKKNLKNRSDVP
YLCHNPAGTSDQLHWAQLKNSQWNGSPLOKMLHYNKQTPPYVDVSAITVPI
AVVNGCHDLADPDQVAMLEPLKPLNLLYKELLPYNHLDFIWMADAPQEVYNEIVTM
AED"
BASE COUNT      398 a 289 c 283 g 380 t
ORIGIN

```

```

Query Match      36.3%; Score 460.2; DB 192; Length 1350;
Best Local Similarity 63.7%; Pred. No. 9.2e-129;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 125 acccagaagcattcatgaattagtgaaatcatccaacatcaaggctatccctgtgag 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 132 ACCCTGAAGCAACATGAATGTAGTCAGATGATACTTACTGGGGATATCCAAGTGAGG 191

QY 185 aatatgaagtcgaactgaagtggtgtatatctcttctgttaacagagattccctgggcc 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 192 AATATGAAGTGTACTGAAGATGGCTACATCTCTGGGGGTCTATAGAATTCCTTATGGGA 251

QY 245 tagtgcacctaagaagacaggttcagccgtgtgtgttactgcagcagtcgctagttg 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 252 ARAAAATTCGTGAGATATCGGCAGAGACCTGTGGCATATTTGACACATGGTTGATG 311

QY 305 gaggtgttagcaactgatttccaacctgcccaacataagcctgggtgttattcttcaggag 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 312 CATCAGGCACAACTGCATTAACAAATCTGCCAAACACAGCCTGGCCTTCATTCAGCAG 371

QY 365 atgctgttttgacgtgtgatggggaacagcagggggaacgcctgggtctcgaaacacaa 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 372 ATGCTGGCTATGATGTGTGCTGGGGAACAGTCGAGGGAATACATGTGTCGGGAAATG 431

QY 425 agacactctcatagaccagagatgagttctgggttttcagttatgatgagatggctaggt 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 432 TATACTATTACCAGACTCAGTTGAATTCGGGCTTTCAGCTTTGATGAATGGCTAAAT 491

QY 485 ttgaccttctgcagtgataaaattttttgcagaaacggccaggaagaaagatctatt 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 492 ATGACCTTCAGGCACCATAGACTTTCATTGTACAGAAACTGGACAGAGAGATACACT 551

QY 545 atgctgcgtattcacagggcacacacatgggtttattgtcattttccaccatgccagagc 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 552 ATGTGTGCTACTTCAGGGCACCACATATCGGTTTATTTGCTTTCTPAACCAATCCTGCTC 611

QY 605 tggctcagaatacaaatgtatttcttagcaccatagcagcagctgttaagcatgcaa 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 612 TGGCTAAANAAYCNACAGGTTTATGTCATAGCTCCAAATGCTGCTGTGAGTATACAG 671

QY 665 aaagccccgggaccacaaatttttggctgcagatatgatgatgaagggattgtttggca 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 672 AAGTCCCTTAAAGACATTTCACTTATTCCTAAGTTTCTCACAAGGTATTTGTA 731

QY 725 aaaaagaatttctgtatcacagcagatttctcagacacactgttattacatttgggcc 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 732 ACAAAATGTTTCATCCGCCACAACTACTTATGATCAATTTCTTGTGACGAAAGTGTCTAC 791

QY 785 agtgattcttgacagattgttagtaataatcatgttacttctgggtgggttaccacaa 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 792 GGGAGCTGTAGATCTCTCTGACGACCAACGCTTTATTCTCTCTGTGGGATTGACAAAGA. 851

```

```

QY 845 acaatatgaacatgagccgagcaagtgtatatgtgccccacactcttgcgtggaacatctg 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 852 AAAAATTAAATGTGAGTCGCTTTGTATCTGTATCTAGGCATATATCCAGCAGGAACATCTA 911

QY 905 tgcataatattctacactgagccagcagtcagtgaaattctgtgtaactcgggcatctgact 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 912 CTCAAGACCTTTTCCACTGGGCACAGCTTGTAAATCTGGGAAGCTTCAAGCCCTATAACT 971

QY 965 gggggagtgagacacaaatcttgaaaaatgcaatcagcaccacccctctgtaaggtacagag 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 972 GGGGAAGTCCATTACAGAACTGTTACACTACATACAGAAACCCCTCCCTACTATGATG 1031

QY 1025 tcagagatatgacggtccctcacagcaatgtggacagaggtcaggaactgggttccaaatc 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1032 TGTGAGCCATGACCGTGCCCAATTCAGTGTGGAAACGGTGGCCATGACATCCCTGGCTGATC 1091

QY 1085 cagaagacgtgaaaaatgctctctctgtgagtgaccacacctcatctacataagaatattc 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1092 CCCAAGATGTGCAATGCTGCTTCCCAACCTCCCAACCTTTTGTACCATTAAGAGATTC 1151

QY 1145 ctgaatggggtccagtggtattctatctgggttgggtgttctcctcacccgtatgtacaatg 1204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1152 TTCCCTACATACCTGGACTTTCATCTGGCGATGGATGCCCTCAAGAGGTTTACAAATG 1211

QY 1205 aatatccatccatctgtatg 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1212 AGATAGTTACCATGATG 1228

RESULT 9
AK010124 1340 bp mRNA HTC 08-FEB-2001
LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DEFINITION library, clone:2310069D09, full insert sequence.
ACCESSION AK010124
VERSION AK010124.1 GI:12845346
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone:2310069D09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Seiuognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1751-1771 (2000)
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1340)

```



```
/db_xref="MGI:1898390"
/db_xref="MGI:1914967"
/clone="2310040L03"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
33.1240
CDS
  /note="putative"
  /codon_start=1
  /protein_id="BAB26466.1"
  /db_xref="GI:12844703"
  /translation="MWLLVTSVLSAFGCAHGLFGKLPKPNENMNVNQMITWGY
SEEYEVVDEGIILGVPIYKKNSENIGRPVAYLOHGLIASINNTITLNPNSLA
FILADAGYVWLGSRNTWGRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQK
GOEKTHYVHSGQITGFIATFNPALAKIKRPVAPVATVRYTESPKKISLIPK
FLVKYIFGNKFMFPHNYLDQFLGTEVCSRELLDLICSNALFICGFDKKNLVSRFDV
VLGNHPACTSTQDILFHWALAKSKLOAYNGSPLONNLHYNQKTPPYDYSAITVPI
AVWNGHIDILADPDQDVAMLLPLKPNLLYKHELPIYNHLDFIWMADAPQEVYNIYTM
AED"
BASE COUNT      395 a      288 c      283 g      380 t
ORIGIN
```

```
Query Match      36.1%; Score 458.6; DB 192; Length 1346;
Best Local Similarity 63.6%; Pred. No. 2.8e-128;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 125 accagaagcattcatgaattagtgaaatcatccaacatcaagcgtatccctgtgag 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 132 ACCCTGAAGCAACATGAATGTTAGTCAGATGATTAACCTTACTGGGGATATCCAAGTGG 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 185 aatatgaagtcgaactgaagtggttcaggccctgtggtgttactgcagcagtggtcctag 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 192 AATATGAAGTGTACTGAAGTGGCTACATCTTGGGGGTCTATAGAATCCTTATGGGA 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 245 tsgtgcacactaagaagacaggttcaggccctgtggtgttactgcagcagtggtcctag 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 252 AGAAAAATCTGAGAAATATCGGCAAGAGACCTGTGGCATATTTGCAGCATGGTTGATTG 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 305 gagggtctgaactgagtttcaacctgcccacaataagcgttcctcattcttcggcag 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 312 CATCAGCCCAAACTGGATTTACAAATCTGCCAAACAAACGCTGGCCTTCATCTCAGCAG 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 365 atgtgtgtttgacgtgtgtgaggggaacagcaggggaaacgcctggtctcgaaacaca 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 372 ATGCTGGCTATGATGTGTGCTGGGGAACAGTCGAGGGAATACATGGTCCCGGAAAAATG 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 425 agacactctccatagaccagaatgagttctgtggtcttcagttatgatgagatggttag 484
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 432 TATACTATTTCACCAAGACTCAGTTGAATCTGGGCTTTTACGTTTGTATGAATGGCTAAT 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 485 ttgacctctctcagtgataaaatttttttcagaaaaacggccaggaaaaagatctatt 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 492 ATGACCTTCAGCCACCACATAGACTTCTATGTGTACAGAAACTGGACAAGAGAGATACACT 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 545 atgtgggtctattcacagggccaccacacctgggtctttattgtcattttccaccatgcccag 604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 552 ATGTTGGTCACTCTCAGGCGACCACTATCGGTTTTATTGSCCTTTTCTACCAATCCTGCTC 611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 605 tggctcagaaaaatacaaaatgtattttgttagcaccctagccactgttaagcagatgcaa 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 612 TGGCTAAAAAATCAAGAGGTTTTATGTCATTAGCTCCAGTTGCTACTGTGAAGTATACAG 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 665 aaagcccccgggaccaaaattttgtgtcgcagatgatgatcaaggatgttttgca 724
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 672 AAAGTCCCTTTAAAAGATTTTCACTTATTCCTAAAGTTTCTTCTCAAGGTGATATTGGTA 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 725 aaaaagaatttctgtatcacagaccagatttctcagacaaactgtttattacctttgtggcc 784
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 732 ACAAAATGTTTCATGCCCAACAACTACTAGATCAATTTCTTGGTACGGGAAGTGTGCTCAC 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 785 aggtgattcttgatcagattgttagtaataatcatgttacttctgtggtgattcaaacca 844
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: August 24, 2001, 21:23:06
Job time: 5573 sec

```
Db 792 GGGAGCTGCTAGATCTTCTCTGCAGCAGCGCTTTATTTCATCTTCTGTGATTGACAAGA 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 845 acantatgaacatgagccgagcaagtgtatatgtctgccccacactcttctggaacatctg 904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 852 AAAACTTAATGTGAGTCGCTTTGTATGTATCTAGGGCATATTCAGCAGGAACATCTA 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 905 tgcnaaatattctacacatggagccaggcagtgaaattctgtgtaactccgggcaattgact 964
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 912 CTCAGAGACTTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGCCTATAACT 971
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 965 ggggagtgagaccnaaaatctggaaaaatgcaaatcagccaactcctctgtaaggtagagag 1024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 972 GGGGAAGTCCATTACAGAACATGTTACACTACAAATCAGAAAAACGCTCCTACTATGATG 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1025 tcaagagatatgacgtccctacagcaaatgtgacacagaggtcaggaactggctttcaaatc 1084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1032 TGTACCCATGACCGTGCCCAATTTGCAGTGTGGAACGGTGGCCATGACATCCTGGCTGATC 1091
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1085 cagagacgtgaaaaatgctctctctgtgaggtgacccaacctcatctaccataagaatttc 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1092 CCCAAGATGTCGAATGCTGCTTCCCANAACCTCCCAACCTTCTGTACCATAGGAGATTC 1151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1145 ctgaatgggctcacgtggaatttcattcattctgtgggttttgatgctcctcaccgtatgtacaatg 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1152 TTCCCTACAATCACCTGGACTTTCATCTGGGCGATGATGCGCTCAAGAGGTTTACAATG 1211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1205 aaatcatccatctctgatg 1221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1212 AGATAGTTACCATGATG 1228
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucelc - nucleic search, using sw model

Run on: August 24, 2001, 20:38:19 ; Search time 211.8 seconds
(without alignments)
6059.636 Million cell updates/sec

Title: US-09-333-159-45

Perfect score: 2044

Sequence: 1 gtcaccacgcgtccggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/NA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/NA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/NA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/NA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/NA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/NA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/NA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/NA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/NA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/NA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/NA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/NA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/NA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/NA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/NA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/NA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/NA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/NA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/NA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/NA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/NA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	22 AAF45131	Human TANGO 294 CD
2	1269	62.1	1269	22 AAF45132	Human TANGO 294 OR
3	489.2	23.9	1378	14 AAO42310	RGL precursor. Or
4	470	23.0	1367	7 AAN60685	Sequence encoding
5	470	23.0	1367	7 AAN60566	Sequence encoding
6	470	23.0	1367	17 AAT58916	Human gastric lipa
7	464.4	22.7	1336	6 AAN50385	Rat prelingual lip
8	444.2	21.7	1528	15 AAO68388	Canine gastric lip
9	444.2	21.7	1528	17 AAT58915	Dog gastric lipase
10	444.2	21.7	1531	15 AAO68389	Canine gastric lip
11	286	14.0	699	22 AAF28679	Human protein HP03

12	286	14.0	1308	22 AAF28689	Human protein HP03
13	257.8	12.6	617	21 AAA44349	Human secreted exp
14	193.2	9.5	572	21 AAA43744	Mouse secreted exp
15	127.6	6.2	936	22 AAF58252	Oligonucleotide D1
16	127.6	6.2	936	22 AAF58254	Oligonucleotide D1
17	127.6	6.2	936	22 AAF58257	Oligonucleotide D1
18	127.6	6.2	936	22 AAF58259	Oligonucleotide D2
19	127.6	6.2	936	22 AAF58262	Oligonucleotide D2
20	127.6	6.2	936	22 AAF58255	Oligonucleotide D1
21	122	6.0	936	22 AAF58252	Oligonucleotide D1
22	122	6.0	936	22 AAF58254	Oligonucleotide D1
23	122	6.0	936	22 AAF58257	Oligonucleotide D1
24	122	6.0	936	22 AAF58259	Oligonucleotide D2
25	122	6.0	936	22 AAF58262	Oligonucleotide D2
26	122	6.0	936	22 AAF58255	Oligonucleotide D1
27	100	4.9	1718	21 AA234958	Soybean acid triac
28	89.4	4.4	1483	21 AA234956	Rice acid triacylg
29	84.8	4.1	859	21 AA234953	Corn acid triacylg
30	72	3.5	1438	21 AA234959	Soybean acid triac
31	58.6	2.9	244	22 AAF58238	Oligonucleotide D1
32	58	2.8	244	22 AAF58238	Oligonucleotide D1
33	55.8	2.7	705	21 AA234952	Catalpa acid triac
34	54	2.6	549	21 AAC93827	Cat flea hindgut a
35	51	2.5	509	21 AA234954	Corn acid triacylg
36	48.2	2.4	178	19 AAV35431	Cell death regulat
37	46.4	2.3	43676	21 AA260606	Nucleotide sequenc
38	42	2.1	8041	17 AAT10105	Adherence conferri
39	40.4	2.0	1140	20 AAX04759	DNA encoding a coc
40	40.4	2.0	1857	21 AAC59142	Human secreted pro
41	40.4	2.0	10628	22 AAF57301	P. falciparum FC83
42	40.2	2.0	2021	19 AAV17087	Homo sapiens tubul
43	40	2.0	1937	21 AAF47445	Maceque TANGO 232
44	40	2.0	2019	21 AAF18286	Lung cancer associ
45	39.4	1.9	3189	20 AAX20284	Borrelia burgdorfe

ALIGNMENTS

RESULT 1
AAF45131
ID AAF45131 standard; cDNA; 2044 BP.
XX
AC AAF45131.
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 cDNA.
XX

Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX

XX
PF 24-MAY-2000; 2000WO-US14858.
XX
PR 14-JUN-1999; 99US-0333159.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX WPI: 2001-032313/04.
XX
DR P-PSDB; AAB66065.
XX

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
XX disease -

PS Claim 1: Fig 6; 359pp; English.

XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.

SX Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;

Query Match 100.0%; Score 2044; DB 22; Length 2044;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	gtcgaccacgcgctccgggaattgcagcaggaataatgtgaagagttttaaaccac	60
DB	1	gtcgaccacgcgctccgggaattgcagcaggaataatgtgaagagttttaaaccac	60
QY	61	aaattcttcttacttagaattagttgtgtacattggcaggaaaaataaatgcagatgtt	120
DB	61	aaattcttcttacttagaattagttgtgtacattggcaggaaaaataaatgcagatgtt	120
QY	121	ggaccatgttgaaacctgtcaagacagtgagttgtctcacacagaatggaaatgtgc	180
DB	121	ggaccatgttgaaacctgtcaagacagtgagttgtctcacacagaatggaaatgtgc	180
QY	181	ttctgattcttggcggtatatttcacagaaatgtgaattcagtacatatgccaaacta	240
DB	181	ttctgattcttggcggtatatttcacagaaatgtgaattcagtacatatgccaaacta	240
QY	241	aagctgtgaccacagaagcattcatgaaattagtagtaaatcccaacatcaaggctatc	300
DB	241	aagctgtgaccacagaagcattcatgaaattagtagtaaatcccaacatcaaggctatc	300
QY	301	cctgtgaggaataatgaatgcgaactgaagatgggtatattcttctgttaacaggattc	360
DB	301	cctgtgaggaataatgaatgcgaactgaagatgggtatattcttctgttaacaggattc	360
QY	361	ctcagagccttagtgcaaccttaagaacagaggttccaggcgtgtgtgttactgcagcatg	420
DB	361	ctcagagccttagtgcaaccttaagaacagaggttccaggcgtgtgtgttactgcagcatg	420
QY	421	gcttagtgaggtgtgtagcaactggaattcccaacctgcccacaataagcctgggcttca	480
DB	421	gcttagtgaggtgtgtagcaactggaattcccaacctgcccacaataagcctgggcttca	480
QY	481	ttctggcagatgctgtgttttgacgtgtgtgaggggaacagcaggggaacgctgtcttc	540
DB	481	ttctggcagatgctgtgttttgacgtgtgtgaggggaacagcaggggaacgctgtcttc	540
QY	541	gaaacacacagacactctccatagaccagaatgagttctgggctttcagttatgatgaga	600
DB	541	gaaacacacagacactctccatagaccagaatgagttctgggctttcagttatgatgaga	600
QY	601	tggctaggttttgacctctcgtgataaaactttattttgcagaaaaacggccaggaaa	660
DB	601	tggctaggttttgacctctcgtgataaaactttattttgcagaaaaacggccaggaaa	660
QY	661	agatctattatgtcgcgtattcacaggccacacacatgggctttatgtcatttccacca	720
DB	661	agatctattatgtcgcgtattcacaggccacacacatgggctttatgtcatttccacca	720

DB	661	agatctattatgtcgcgtattcacaggccacacacatgggctttatgtcatttccacca	720
QY	721	tgccagagctgctcagaaaaataatatttttcttagcaccatagccacagtta	780
DB	721	tgccagagctgctcagaaaaataatatttttcttagcaccatagccacagtta	780
QY	781	agcatgcaaaaaagccccgggaccacaaattttgtgtgcagatatgatcaaggat	840
DB	781	agcatgcaaaaaagccccgggaccacaaattttgtgtgcagatatgatcaaggat	840
QY	841	tgtttggcaaaaaagaatttctgtatcacaccagatttctcagacaaactgttattacc	900
DB	841	tgtttggcaaaaaagaatttctgtatcacaccagatttctcagacaaactgttattacc	900
QY	901	tttttggccagatgattcttgcagatgttagtaataatcatttcttacttctggggat	960
DB	901	tttttggccagatgattcttgcagatgttagtaataatcatttcttacttctggggat	960
QY	961	tcaacacacaaatatgaacatgagccgagcaagtgtatgtgtgccacacactctgtc	1020
DB	961	tcaacacacaaatatgaacatgagccgagcaagtgtatgtgtgccacacactctgtc	1020
QY	1021	gaacatctgtgaaaaatatctacactgagccagcagtgaaattcttggtaactccgg	1080
DB	1021	gaacatctgtgaaaaatatctacactgagccagcagtgaaattcttggtaactccgg	1080
QY	1081	catttgcactggggagtgagacaaaatctgaaaaatctgaaaaatgcaatcagcaactcctgtaa	1140
DB	1081	catttgcactggggagtgagacaaaatctgaaaaatctgaaaaatgcaatcagcaactcctgtaa	1140
QY	1141	ggtacagagtcagagatatgacggtccctacacagcaatgtggacagaggtcagagactggc	1200
DB	1141	ggtacagagtcagagatatgacggtccctacacagcaatgtggacagaggtcagagactggc	1200
QY	1201	tttcaaatccagaagcgtgaaatgtcgtctctctgaggtgacacacactcatcaccata	1260
DB	1201	tttcaaatccagaagcgtgaaatgtcgtctctctgaggtgacacacactcatcaccata	1260
QY	1261	agaaattctcctgaatgggtccagtgatttctcctgggtttggatgctccctcacogta	1320
DB	1261	agaaattctcctgaatgggtccagtgatttctcctgggtttggatgctccctcacogta	1320
QY	1321	gttacaatgaaatcatccatctgtagtcagcagagagagacacacacttcccagggacgt	1380
DB	1321	gttacaatgaaatcatccatctgtagtcagcagagagagacacacacttcccagggacgt	1380
QY	1381	gtgagccgtattgtgaagcatctgacactgacatcttaggacaaacctctcaggagatg	1440
DB	1381	gtgagccgtattgtgaagcatctgacactgacatcttaggacaaacctctcaggagatg	1440
QY	1441	gggctagggaccatgaaggcagaattacagagagcagacactagatatacttttcaga	1500
DB	1441	gggctagggaccatgaaggcagaattacagagagcagacactagatatacttttcaga	1500
QY	1501	ttcccctggcacttggcactaaatccgacacttacatttattttttctgtgaaataaa	1560
DB	1501	ttcccctggcacttggcactaaatccgacacttacatttattttttctgtgaaataaa	1560
QY	1561	gtacttattaggtaaatagaggttttgtatgctattatattctaccatcttgaagggt	1620
DB	1561	gtacttattaggtaaatagaggttttgtatgctattatattctaccatcttgaagggt	1620
QY	1621	aggttttaccctgatagccagaaaaatctcagacattctctatcattcaggttaattctc	1680
DB	1621	aggttttaccctgatagccagaaaaatctcagacattctctatcattcaggttaattctc	1680
QY	1681	tttanaaacaccttattgttttttctataagccatatttttggagcactaaagtaaatggc	1740
DB	1681	tttanaaacaccttattgttttttctataagccatatttttggagcactaaagtaaatggc	1740
QY	1741	aaatttgggacagatatggaggtctgtgagctgtggattattgttgcatttgacaaaaata	1800
DB	1741	aaatttgggacagatatggaggtctgtgagctgtggattattgttgcatttgacaaaaata	1800

QY 1801 gctagacattttcaacctgtgtgcccacagagacataaacactacctcaggaagctgagctgc 1860
 Db 1801 gctagacattttcaacctgtgtgcccacagagacataaacactacctcaggaagctgagctgc 1860
 QY 1861 tttaaggacaacaacaacaatcagttgttacagtagatgaaatctatttaagcatt 1920
 Db 1861 tttaaggacaacaacaacaatcagttgttacagtagatgaaatctatttaagcatt 1920
 QY 1921 ctcaagaataaggccaagttttatagttcattctcagggaagaaaaattttataggaatttt 1980
 Db 1921 ctcaagaataaggccaagttttatagttcattctcagggaagaaaaattttataggaatttt 1980
 QY 1981 atgaatttccaataaaatgcattctgcattacataaaaaaataaaaaaaggcg 2040
 Db 1981 atgaatttccaataaaatgcattctgcattacataaaaaaataaaaaaaggcg 2040
 QY 2041 ccgc 2044
 Db 2041 ccgc 2044

RESULT 2

AAF45132
 ID AAF45132 standard; CDNA; 1269 BP.

XX AAF45132;

DT 30-MAR-2001 (first entry)

XX Human TANGO 294 ORF.

DE Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

XX Homo sapiens.

XX WO200077239-A2.

PD 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

DR P-PSDB; AAB66065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease.

XX Claim 1; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, and
 CC autonomic function disorders such as hypertension and sleep disorders,

CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence: 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

Query Match 62.1%; Score 1269; DP:22; Length 1269;

Best Local Similarity 100.0%; Pred.No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 atgttggaaaacctgtccaagacagtggtgtctcacacagaaatgaaatgggcttcg 195

Db 1 atgttggaaaacctgtccaagacagtggtgtctcacacagaaatgaaatgggcttcg 60

QY 186 atctctgggtggcgatatgttccagagaaaatgtgaattcagtagacatgccaactaaagct 245

Db 61 atctctgggtggcgatatgttccagagaaaatgtgaattcagtagacatgccaactaaagct 120

QY 246 gtggaccacagaagcattcatgaattatgtgaaaatccaacatacaagcctatccctgt 305

Db 121 gtggaccacagaagcattcatgaattatgtgaaaatccaacatacaagcctatccctgt 180

QY 306 gaggaatatgaagtcgcaactgaagatgggtatatctcttctgtaaacaggattccctga 365

Db 181 gaggaatatgaagtcgcaactgaagatgggtatatctcttctgtaaacaggattccctga 240

QY 366 ggcctagtgcaacctaaagaacaggttccagggcctgtggttactgcagcatggccta 425

Db 241 ggcctagtgcaacctaaagaacaggttccagggcctgtggttactgcagcatggccta 300

QY 426 gtggaggtgttagcaactggaatttccaacctgccacaataagcctgggcttcatctcg 485

Db 301 gtggaggtgttagcaactggaatttccaacctgccacaataagcctgggcttcatctcg 360

QY 486 cgcagatcgtgtgttgacgtgtggtgggaacacagcaggggaaacgcctggctcga 545

Db 361 cgcagatcgtgtgttgacgtgtggtgggaacacagcaggggaaacgcctggctcga 420

QY 546 cacaagacacctctccatagaccaagatgagttctcgggttccagttatgatagatggct 605

Db 421 cacaagacacctctccatagaccaagatgagttctcgggttccagttatgatagatggct 480

QY 606 aggtttgacctctcgcagtgataaaccttatttgcagaaaacggggccagaaaagatc 665

Db 481 aggtttgacctctcgcagtgataaaccttatttgcagaaaacggggccagaaaagatc 540

QY 666 tattatgtcggctattccacagggccaccacatggccttattgcattttccaccatgccca 725

Db 541 tattatgtcggctattccacagggccaccacatggccttattgcattttccaccatgccca 600

QY 726 gagctggctcagaaaaatcaaaatgtatttggcttttagcaccatagccactgttaagcat 785

Db 601 gagctggctcagaaaaatcaaaatgtatttggcttttagcaccatagccactgttaagcat 660

QY 786 gcaaaaaagcccggaacaaaattttgtgtgcagatatgatgatcaagggatgttt 845

Db 661 gcaaaaaagcccggaacaaaattttgtgtgcagatatgatgatcaagggatgttt 720

QY 846 ggcaaaaaagaattctctatcagaccagatttccagacaactgtttattacccttgt 905

Db 721 ggcaaaaaagaattctctatcagaccagatttccagacaactgtttattacccttgt 780

QY 906 ggccagggtgattcttgatcagattttagtaataatcatgttacttcttggtggatccaac 965

Db 781 ggccagggtgattcttgatcagattttagtaataatcatgttacttcttggtggatccaac 840

QY 966 accacataatgaacatgacgcagcaagtgatatgtcgcacacactcttctgtgaaaca 1025

Db 841 accacataatgaacatgacgcagcaagtgatatgtcgcacacactcttctgtgaaaca 900

QY 1026 tctgtgcaaaatattctacactggacgcagcaggaattctgtgaaacctggcgcat 1085

Db 901 tctgtgcaaaatattctacactggacgcagcaggaattctgtgaaacctggcgcat 960

QY 1086 gactgggggagtgagaccacaaatcttgaaaaatgcaatcagccaaactctctgtaaggtac 1145
DB 961 gactgggggagtgagaccacaaatcttgaaaaatgcaatcagccaaactctgtaaggtac 1020
QY 1146 agagtcagagatatgacggtccctcacagcaaatgtggacagaggtcaggaactggttca 1205
DB 1021 agagtcagagatatgacggtccctcacagcaaatgtggacagaggtcaggaactggttca 1080
QY 1206 aatccagaagacgtgaaaaatgctctctgaggtgaccaaactctctaccataaagaat 1265
DB 1081 aatccagaagacgtgaaaaatgctctctgaggtgaccaaactctctaccataaagaat 1140
QY 1266 attcctgaatgggctcagtggaattcatctatctggggttggatgctctccacgtatgtac 1325
DB 1141 attcctgaatgggctcagtggaattcatctggttggatgctctccacgtatgtac 1200
QY 1326 aatgaatatccatctgacgagcagagagagaccaaactcttccaggaagcgtgtgag 1385
DB 1201 aatgaatatccatctgacgagcagagagagaccaaactcttccaggaagcgtgtgag 1260
QY 1386 gccgtattg 1394
DB 1261 gccgtattg 1269

RESULT 3

AAQ42310
ID AAQ42310 standard; DNA; 1378 BP.

AC AAQ42310;

XX 20-SEP-1993 (first entry)

DE RGL precursor.

XX Rabbit gastric lipase; RGL; PRGLN2.1; fat; bioconversion;
hydrolysis; transesterification; ds.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers
FH mlec_RNA 120..1247

FT /tag= a

FT /note= "claim 1; page 10-11"

XX EP543629-A.

XX 19-MAY-1993.

XX 12-NOV-1992; 92EP-0403055.

XX 13-NOV-1991; 91FR-0013948.

XX (LJOU) INST RECH JOUVEINAL.

XX Benicourt C, Blanchard C, Junien J;

XX WPI; 1993-161080/20.

XX P-PSDB; AAR37302.

XX Rabbit gastric lipase, its precursor and their DNA - useful for
treating conditions linked to gastric lipase deficiency, such as
mucoviscidiosis and pancreatic exocrine insufficiency

XX Claim 1; Fig 6; 31pp; French.

XX mRNA was isolated from rabbit gastric mucosal cells, converted to
cDNA and fragments cloned in pUC18. Recombinant plasmids were used
to transform E.coli MM294 and cells screened for reactivity with
CC probes corresp. to known parts of the RGL sequence. One positive
clone contd. plasmid pJ0101 which contd. a 1.35 kb PSTI sequence.
CC This fragment was isolated, subjected to PCR and the amplification

CC prod. ligated with the 2.2 kb BglII-SacI fragment of pRU276 (contg.
CC a synthetic Tac promoter and transcription stop signal). The
CC recombinant plasmid was designated PRGLN2.1.
XX
SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;

Query Match 23.9%; Score 489.2; DB 14; Length 1378;

Best Local Similarity 65.0%; Pred. No. 1.5e-115;

Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 250 acccagaagcattcatgaattattagtgaaatcatccaaatcaggtctatccctgtgagg 309
DB 136 accctgaagtgaatgaattattagtcagatgattctctactggggaaccgaagtga 195
QY 310 aatatgaagtcgaactgaagatgggtatatactcttctgttaacagagattccctcgaagcc 369
DB 196 aatatgaagtcgtgactgaagatgggtatatactcttctgttaacagagattccctcgaagcc 255
QY 370 tagtgcacactgaagacaggttccaggtccaggtggtggttactgcagcagcgttagttg 429
DB 256 agaaaaatttcagggaacagagagccgctgatttttgcagcagcgttgggtgcttg 315
QY 430 gagggtcagcaactggatttccaaactgcacaaatagcctggggtcttctattctgtgcag 489
DB 316 catcagcctcaaaactggatctccaaactgcacaaacagcgtgcttattcttgcag 375
QY 490 atgctggttttgacgtggtgagtggaacagcagggggaacgcgtgctctcgaacacaca 549
DB 376 atgctggttctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 435
QY 550 agacactctccatgacacaaatgagttctggtggttcttgcagtaattgagatgagatgaggt 609
DB 436 tatactattccacagactcgttgaaattctggggttcttgcagtaattgagatgagatgaggt 495
QY 610 ttgaccttctgagtgataaaactttatttgcagaaacggccaggaagaaatctatt 669
DB 496 atgaccttccagcccaattgacttcttgcagaaacggccaggaagaaatctatt 555
QY 670 atgctggtattccacagggcaccacacacacacacacacacacacacacacacacacacac 729
DB 556 atgctggtattccacagggcaccacacacacacacacacacacacacacacacacacacac 615
QY 730 tggctcagaaaaatcaaaatgattttgctttagcaccacacacacacacacacacacacacac 789
DB 616 tggctcagaaaaatcaaaac 675
QY 790 aaagcccccgggacacaaatttttgcctgcagatgatatgacagggggtggttggca 849
DB 676 aaagccttgtaacaaacttaggtttattctccaaactggttcaagattatttgggtg 735
QY 850 aaaaagaatttctgatacagacagacacacacacacacacacacacacacacacacacacac 909
DB 736 acaaaatattctaccacacaaatttcttgcacaaatttcttgcacacacacacacacacacac 795
QY 910 aggtgatttctgacagatttgtagtaataatcatgttacttctgggtgggttcaacacaca 969
DB 796 gtgagacactgaatgaatttgcagcaatgcttatttatttgcagcagcgttgcagcagc 855
QY 970 acaatatgaactgagccgagcgaagtgtatgtctgccacacacacacacacacacacacacac 1029
DB 856 caaaattgacacatgagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 915
QY 1030 tgcacaaattctacactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1089
DB 916 ttcacaaactgctgcactggaccccggtgtttaaactgggaatttcaagcgttcaatt 975
QY 1090 ggggggagtgagacacaaatctggaataatgcaatcagccaaactctctgtaaggtacagag 1149
DB 976 gggggaagtcagctcagaatgtagtgcacttcaatcagccacacacacacacacacacacacac 1035
QY 1150 tcagagatgacggttccctacagcaatgtggacagaggtcaggaagcgtggttcaaatc 1209

Db 1036 tgaccgcccataaattgccaattgcagctgtggagtggtggcaatgactggtggtgacct 1095
Qy 1210 cagaagacgtgaaaatgctctctgagtgagtgacacacccatctacataagaatattc 1269
Db 1096 cccaagatgtgaccttttgcctccaaaacttctcaatctatttaccacagagattc 1155
Qy 1270 ctgaatgggtcagctgatttctctggttttggtggtgctcctcaccgtatgtacaatg 1329
Db 1156 ttccataatacacttgattttatcttggtggaatgactcctcacaagaatttacaatg 1215
Qy 1330 aaatcatccatctgagcagcagagagaga 1359
Db 1216 aaattattctatgatggcaaaagataaaa 1245

RESULT 4

AAN60685

ID AAN60685 standard; DNA; 1367 BP.

XX

AC AAN60685;

XX

DT 08-JUN-1991 (first entry)

XX

DE Sequence encoding pregastric lipase.

XX

KW Precursor polypeptide; secretion vector; enzyme; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT sig_peptide 47..103

FT /*tag= a

FT mat_peptide 104..1243

FT /*tag= b

XX

PN W08603778-A.

XX

PD 03-JUL-1986.

XX

PF 23-DEC-1985; 85WO-GB00599.

XX

PR 21-DEC-1984; 84GB-0032483.

XX

PR 23-DEC-1985; 85WO-GB00599.

XX

PR 01-JAN-1986; 86GB-0019568.

XX

XX (BREW-) BREWING RES FOUND.

PA (TUBB/) TUBB R S.

XX

PI Tubb RS;

XX

XX WPI; 1986-182910/28.

XX

DR P-PSDB; AAP60724.

XX

PT New precursor polypeptide of defined sequence - and corresp. DNA

XX

PT used to transform hosts for prodn. of the polypeptide

XX

PS Example; Fig 5; 60pp; English.

XX

CC The inventors claim a DNA sequence encoding AAP60723 linked to a

XX

CC promoter upstream and a gene for a polypeptide downstream.

XX

CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,

XX

CC gastric lipase and the mammalian lymphokine, interferon-alpha2.

XX

SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

Query Match

Best Local Similarity 23.0%; Score 470; DB 7; Length 1367;

Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

Qy 244 ctgtggaccagacattcatgaattattagtgaaatcatccaaatcaaggctatccct 303

Db 123 ctggagccctgaagtgactgaattgaacattagtcagatgatttcttggtggataccaa 182

Qy 304 gtcaggaaatgaagtcgcaactgaagatgggtatatactcttctgttaacaggattccctc 363
Db 183 atcagaataatgaagatgtgactgaagatgttatctcttgagatcaatagaattccctt 242
Qy 364 gacgctagtgcacactaagaagacaggttccagagcctgtggtgtactgcagatggcc 423
Db 243 atcgaagaataatcaggaataacagccagagacgtgtgtgttttgcagcatgggt 302
Qy 424 tagtgagggtgctagcaactggatttccaaactctcccaacaaatgagcctgggtctcacc 483
Db 303 tgcctgcatcagccacaaactggatttccaaactctcccaacaaactgctctcacc 362
Qy 484 tggcagatgctggttttgacgtgtggatggggaacagcaggggaacgctggtctcgaa 543
Db 363 tggcagatgctggttttatgatgtgtgtgtggcgaacagcagagaaacacacgtggcagaa 422
Qy 544 aacacagacactctccatagacaaagatgagttgtggttcttggttctagttatgatgatg 603
Db 423 gaaactgtactattccacagattcagttgaattctgtggttctgagcttttgatgaatgg 482
Qy 604 ctgagtttgactctctcagtgatgataaactttattttgcagaaaacggccaggaaga 663
Db 483 ctaaatagactctccagcccaactcagcttcaattgtaaaagaaactggacagagcagc 542
Qy 664 tctattatgctggtctattccacagggcaccacacacacacacacacacacacacacac 723
Db 543 tacactatgttggccatttccacagggcaccacacacacacacacacacacacacacac 602
Qy 724 cagagctggtcagaaaatacaaaatgtattttgtttgagcaccacacacacacacacac 783
Db 603 ccagctggtcagaaaatacaaaacttctatgctgctgagctctgttgcacatgtgaatg 662
Qy 784 atgcaaaaagcccggggacaaatgttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 843
Db 663 atacaaaagccttataaaacaaacttagattgttctctcaacacacacacacacacacac 722
Qy 844 ttggcaaaaagaatttctgtatcagacacagatttctcagacacacacacacacacacac 903
Db 723 ttgtgacaaaataattctaccacacacacacacacacacacacacacacacacacacac 782
Qy 904 gtggccaggtgattctgtatcagattttagtaataatcatgttaacttctggttggaatca 963
Db 783 gctccgtgagatgctgaattctcttgcagcaatgcttattataatttggatttg 842
Qy 964 acaccaacaataatgaacatgagccgagcaagtgatagctgtgtgtgtgtgtgtgtgtgtgt 1023
Db 843 acagtgaagaactttaacacagagtcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 902
Qy 1024 catctgtcacaataattctacactggagccagcgagtgaaattctgtgtgaactccggcgt 1083
Db 903 ctctgttcaaaaacatgttccattgtgacccagcgttgaagtctgtggaaattccaagctt 962
Qy 1084 ttgactggggagtgagacccaataatctggaaaaatgcaatcagcccaactctctgtgaagt 1143
Db 963 atgactggggagcagcttccagaaatgagtgacatgacatgacatgacatgacatgacatg 1022
Qy 1144 acagagtcagagatatgacggtccctcagacaaatgtgacagagagtgagagtgagagtg 1203
Db 1023 acaatgtgacagccatgaatgtaccaaatgtgagtgagagcgtgtggagagcgtgtgtgtg 1082
Qy 1204 caaatcccaagaagcgtgaaaaatgctgctctctgaggtgacccacacacacacacacac 1263
Db 1083 ctgaccccaagaatgtgtgcttcttccaaaactcccaactcttattattaccacaaagg 1142
Qy 1264 atattcctgaatgggctcagctggatttcaatctgtgggttggatgctctcaccgtatgt 1323
Db 1143 agattccttttacaatcacttggttcttatctgtggcaatggatgctccctcaagaagttt 1202
Qy 1324 acaatgaatcatccatctgat 1345
Db 1203 acaatgacattgttctatgat 1224

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 47..1243
 FT CDS /*tag= a
 FT /product= gastric lipase
 FT sig_peptide 47..103
 FT /*tag= b
 FT mat_peptide 104..1240
 FT /*tag= c
 XX W09633277-A2.
 XX 24-OCT-1996.
 XX 19-APR-1996; 96WO-FR00606.
 XX 20-APR-1995; 95FR-0004754.
 XX (BIOC-) BIOCEM SA.
 XX (LJOU) INST RECH JOUVEINAL.
 XX Baudino S, Benicourt C, Cudrey C, Gruber V, Lense P;
 PI Merot B;
 XX WPI; 1996-485783/48.
 DR P-PSDB; AAM09383.
 XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 FT plants - useful for facilitating absorption of fat, as bio-catalysts
 FT and for prodn. of bio-fuel
 XX Claim 3; Fig 4; 130pp; French.
 XX This is the nucleotide sequence encoding the human pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 CC the mature protein, respectively) can be deleted to form the derivatives
 CC designated delta-4 or delta-54 respectively. Plants, or their extracts,
 CC expressing the lipases or the truncated derivatives, can be used:
 CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat absorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 SQ
 Query Match 23.08; Score 470; DB 17; Length 1367;
 Best Local Similarity 64.28; Pred. No. 1.2e-110;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 QY 244 ctgtggaccagagcattcatgattagtgaaatcaccacacatcaaggtatccct 303
 DB 123 ctggaaacctgaagtactatgaacattagtcagatgattactattggggatcccaa 182
 QY 304 gtgagggaatgaagtcgcactgaagatgggtatattctcttctgttaacaggattctc 363
 DB 183 atgaagaatgaagtgtgactgaagatggttatattcttgagtcgaatgaattcctt 242
 QY 364 gaggcctagtgcacacctaaagaacagagttccaggccctgtggttactcagcagatgcc 423
 DB 243 atgggaagaaataatcagggaatacagccagagacctgtgtgtttttcagcagatggt 302
 QY 424 tagttgaggtgtcagcaatggatttccacacctgcccacaaatagctgggtcttattc 483
 DB 303 tgcctgcatcagcccaactggatttccacacctgcccgaacacagcctgtgccttattc 362

QY 484 tggcagatgtgtgttttggacgtgtggtggtggaacacagcaggggaaacccctgtctogaa 543
 DB 363 tggcagatgtgtgttttggacgtgtggtggtggaacacagcaggggaaacccctgtctogaa 422
 QY 544 aacacaagacactctccatagacacacagatgagttctctgggttctcagttatagatagg 603
 DB 423 gaaacttgactattccaccagattcagttgaaattctctgggttctcagttatagatagg 482
 QY 604 ctgaggttgcctctcctgcagtgataaaactttatttgcagaaacacgagcaggaagaaga 663
 DB 483 ctaaatatgacctccagccac 542
 QY 664 tctattatgtcgtctatccacagggccacacacacacacacacacacacacacacacacac 723
 DB 543 tacactatgtggcattccacagggccacacacacacacacacacacacacacacacacac 602
 QY 724 cagagctggctcagaaatacaaaatgattttgttctttagcaccacacacacacacacac 783
 DB 603 ccagcctggctaaagaatcaaaacaccttctatgctctagctcctgttgcacctgtgaagt 662
 QY 784 atgcacaaagcccccgggacacaaattttgttctgcgcagatattgatcatcaaggatgtg 843
 DB 663 atacaaagaccttataac 722
 QY 844 ttggcaaaaaagaatttctgtatcagaccagatttctcagacacacacacacacacacac 903
 DB 723 ttggtgacaaatatttaccac 782
 QY 904 gtggcaggtgattcttgatcagattgttagtaataatcatttactctcgttggtgagttca 963
 DB 783 gctcccgtagatgctgaattcctcttgcagcaatgcttattataatttggatttg 842
 QY 964 ac 1023
 DB 843 acagtaagaactttaacacagcagctggtggtggtggtggtggtggtggtggtggtggt 902
 QY 1024 catgtgcaaaaaattcttactacactggagccagcagtgaaattcttggtgaaactcgggcat 1083
 DB 903 ctttctgttcaaaacacatttccattggacccagcgtttaaagctcgggaaattcccaagctt 962
 QY 1084 ttgactgggggagtgagacacaaatactggaaaaatgcacacacacacacacacacacac 1143
 DB 963 atgactgggggagccagcttccagaaatgagtgacatgacatgacatgacatgacatgac 1022
 QY 1144 acagatcagagatgacagctccctcagcaatggtgagcagaggtcagactggcctt 1203
 DB 1023 acantgacagcagcagatgaatgtaccacattgacagtggtgagcaggtgagcagcctgtg 1082
 QY 1204 caantccagaaagacgtgaaaaatgctctctctgaggtgagcagcagcagcagcagcag 1263
 DB 1083 ctgaccccaagatgttggccttttcttccaaaactcccaactcttatttaccacaggg 1142
 QY 1264 atattctcgaatgggtcagcagtgatttctcctctctctctctctctctctctctctct 1323
 DB 1143 agattctcttttaacacacttgactttattctgtggcaatggtgacccctcaagaagttt 1202
 QY 1324 acaatgaatcatccatctgat 1345
 DB 1203 acaatgacattgttctatgat 1224

RESULT 7

AAN50385
 ID AAN50385 standard; DNA; 1336 BP.
 XX
 AC AAN50385;
 XX
 DT 17-JAN-1992 (first entry)
 XX
 DE Rat prelingual lipase gene.
 XX
 KW Prelingual lipase; enzyme; EC-3.1.1.3; ss.

626	acaaaattctaccacaccactcttttgatcaattctcgccacgaggtatgctccc	685
QY	agtgattcttgatcagatttgtagtaataatactattacttctgggtggtatcaacca	969
Db	gcgagacggtgggattcccttcgacacgcctgttatcatgttggtatgacta	745
QY	acaatatgaacatgagccgagcaagtgtatatgtgcgccacactcttctgggaacatcg	1029
Db	tgaacttgaacatgagctcgctggatgtgtatctgtcaataatccagcaggaacatcgg	805
QY	tgcuaaataattcacactggagccgagcgatgtaattctgttgaaactcggcgatttgact	1089
Db	ttcagaaactgtctccactggtccagcgtgttaagtctgggaagtccaaacttttgact	865
QY	gggggagtgagacaaaatactggaaaaatgcaatcagcccaactctgtaagggtacagag	1149
Db	gggtgagcccgcttcgagaacatgtagcactatcatcagagcatgctcctactacaacc	925
QY	tcaagatatgacggtccctacagcaaatgtggacgaggggtcaggaactggctttcaaatc	1209
Db	tgacagacatgcatgtgccaatcgcagtggtggaacggtggcaacgactgtgtgcccagcc	985
QY	cagaagacgtgaaaaatgctctctctgtgaggtgaccaaacctcatctaccataagaatttc	1269
Db	ctctcgatgttgacctttgcttccaaactcccaatctcatattaccacagggaattc	1045
QY	ctgcaatgggctcacgtggatttcaatctggggtttggatgctctccacacgtatgataaatg	1329
Db	ctccttacaataacttggacttattctggggccatggaagcccccaacggtttacaatg	1105

1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
20

QY	1130	aaatccatccatcatg	1346
Db	1106	aaattgttccatgatg	1122
RESULT 10			
AAQ68389	standard; DNA; 1531 BP.		
XX	AAQ68389;		
XX	20-FEB-1995	(first entry)	
XX	Canine gastric lipase (Met-CGL).		
XX	Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;		
KW	bioconversion; exocrine pancreatic insufficiency; ss.		
XX	Canis familiaris.		
OS			
XX	Key	Location/Qualifiers	
XX	FT CDS	1..1143	
XX	FT	/*tag= a	
XX	WO9413816-A.		
XX	23-JUN-1994.		
XX	16-DEC-1993;	93WO-FR01260.	
XX	16-DEC-1992;	92FR-0015201.	
XX	(LJOU) INST RECH JOUVEINAL.		
XX	Benicourt C, Blanchard C, Junien J;		
XX	WPI: 1994-217890/26.		
XX	P-PSDB; AAR56871.		
XX	Recombinant canine gastric lipase and nucleic acid encoding it -		
XX	are used for improving absorption of ingested fat, treating		
XX	mucoviscidosis etc. and in enzymatic bio-conversions		
XX			

RESULT 10	
AAQ68389	
ID	AAQ68389 standard; DNA; 1531 BP.
XX	
AC	AAQ68389;
XX	
DT	20-FEB-1995 (first entry)
XX	
DE	Canine gastric lipase (Met-CGL).
XX	
KW	Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
KW	bioconversion; exocrine pancreatic insufficiency; ss.
XX	
OS	Canis familiaris.
XX	
FH	Key
FT	Location/Qualifiers
ET	1..1143
XX	/*tag- a
XX	
PN	W09413816-A.
XX	
XX	
PD	23-JUN-1994.
XX	
PF	16-DEC-1993; 93WO-FR01260.
XX	
PR	16-DEC-1992; 92FR-0015201.
XX	
PA	(LJOU) INST RECH JOUVEINAL.
XX	
PI	Benicourt C, Blanchard C, Junien J;
XX	
XX	WPI; 1994-217890/26.
DR	P-PSDB; AAR56871.
XX	
PT	Recombinant canine gastric lipase and nucleic acid encoding it -
PT	are used for improving absorption of ingested fat, treating
PT	mucoviscidosis etc. and in enzymatic bio-conversions

Claim 2: Fig 8; 52pp; French.

The sequence given below is the sequence of figure 8, altered according to the amendments described on page 2 of the appended letter.

CGL is used to improve absorption of ingested fat, in healthy and sick patients (e.g. having altered levels of gastric lipase); to treat conditions associated with insufficiency (or lack) of lipases, esp. mucoviscidosis or exocrine pancreatic insufficiency and partic. where immobilised, for bioconversions, e.g. hydrolysis or transesterification (other mammalian gastric lipases, or derivs., can be used in this application).

Sequence 1531 BP; 395 A; 386 C; 329 G; 421 T; 0 other;

Query Match 21.7%; Score 444.2; DB 15; Length 1531;
Best Local Similarity 62.8%; Pred. No. 5.1e-104;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 accaagaagcattcatgaattatgtagaatcatccaacatcaaggctatccctgtgagg 309
DB 29 accctgaagtcaccatgaataagtcagatgacactactcctgggataccagctgagg 88
QY 310 aatatgaagtcgaactgaagatgggtatatctcttctgttaacagattccctcgagcc 369
DB 89 aatatgaagttgaccgaagcaggttatctccttggatcgacagaattccttatggga 148
QY 370 tagtgcacctgaagaagacaggttccagcctgtgggttactgcagcatggcctagtctg 429
DB 149 ggaataattcagagaataatagcggagacctgtgtcattttgcaacacggttgctcg 208
QY 430 gaggtgctagaactggatttccaacctgccacaatagcctggccttccatctctggcag 489
DB 209 catcagccacaactggatctccaacctgccacaacagcctggcctcctcctggccg 268
QY 490 atgctggtttgacgtgtggtatggggaacagcaggggaacgcctggtctcgaatacaca 549
DB 269 acgcgggtacagctgtggctgggggaacagcaggggaacacctgggcaggaggaaatc 328
QY 550 agacactctccatgaccaaagatgagttctgggctttcagttatgatgatgagtgctaggt 609
DB 329 tgtactactgcgccgactccgtcgaaattctggctttcagctttgacgagatggtcaaat 388
QY 610 ttgacctctctcagtgataaacatttatttgcgaaaaacggccagggaaaagatctatt 669
DB 389 atgacctctccgccaccattgacttcatttgaagaaacgggacagggacaagctacat 448
QY 670 atgtcggtctatcacagggaccaccatggccttatttgcattttccaccatgccagacg 729
DB 449 acgttggccattccagggccaccaccttggttctatgcgtttccaccacaaccacagc 508
QY 730 tggctcagaaaaataatgtatttggcttagcaccctatagcactgtttaagcatgcga 789
DB 509 tggcgaaacgcatcaaaacctctatgatctagctccgttgcacogtgaagtacacgcg 568
QY 790 aaagcccccgggaccaaaattttgtgtgcgagatatgatcataagggattgtttggcca 849
DB 569 aaaccctgttaaacaaactcatgctgcctctcgttccctcctcaagcttatatttggaa 628
QY 850 aaaaagaattctgatcagaccagatttctcagacaactgttatttaccctttgtggcc 909
DB 629 acaaaatattctaccacaacactcttctgataatttctgcacacgagtgatgtccc 688
QY 910 aggtgattcttgatcagattttgtagtaataatcatttactctgtgttgattcaacacca 969
DB 689 gcgagacggttgatctcctctgcagcaacgccctgtttatcatttgggtatttgacacta 748
QY 970 acaatatgaacatgacgagcaagtgatatgtgccacactcttctgtcgaaactctg 1029
DB 749 tgaacttgaaatgagtcgcttggaatgagttatctgtcacaataatccagcaggaacatcgg 808

CC utility in compositions used for bone, cartilage, tendon and/or nerve
CC tissue growth or regeneration as well as wound healing and in the
CC treatment of burns. The proteins may be used in the treatment of
CC periodontal disease and in other tooth repair processes. Other uses
CC include treatment of thrombolytic and haemostatic conditions, treatment
CC or prevention of tumours and inhibiting infection by bacteria, viruses,
CC fungi and other parasites.
XX
SQ Sequence 699 BP; 216 A; 144 C; 148 G; 191 T; 0 other;

Query Match 14.0%; Score 286; DB 22; Length 699;
Best Local Similarity 67.4%; Pred. No. 1.2e-63;
Matches 403; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 211 gaaatgtgaattcagtcacatgccaaactaaagctgtggaccacgaagcattcatgaata 270
DB 44 gatctatgtgttatgacaaagaagaacaaatgcaaaccttgagctaaatgaata 103
QY 271 ttagtgaatccacacatcaagggtatccctgtgaggaaatgaagtcgcaactgaag 330
DB 104 ttaggcagattattcttactgggttattcttataagagagatgattttacacaaaag 163
QY 331 atgggtatattcttctgttaacagagattctctgagccttagtgcaacctaaagaacag 390
DB 164 atgggtatattcttctgttaacagagattctctgagccttagtgcaacctaaagaacag 223
QY 391 gtccagcctgtgttactcagcagctgcttagtgagagctgagcagcagcagcagcagc 450
DB 224 ctcaaacgctgtgttattcagcagctgcttagtgagagctgagcagcagcagcagcagc 283
QY 451 ccaacctgcccaacatagcctggcttcttctgagagatgctgtgttgacgtgtgga 510
DB 284 gcaacctgcccaacatagcctggcttcttctgagagatgctgtgttgacgtgtgga 343
QY 511 tggggacacagcaggggaacacgctgtgtctgaaacacacagcagcagcagcagcagcag 570
DB 344 tggggacacagcaggggaacacgctgtgtctgaaacacacacgctgtgtgtgtgtgtgt 403
QY 571 atgagctgt 630
DB 404 cagaatactggcctcagttgt 463
QY 631 actttatttgcagaaacagggccaggaagaagatctattatgtcgtgtgtgtgtgtgtgtgt 690
DB 464 attttatcagagaacacggtgacagagcgactctactacgtgtgtgtgtgtgtgtgtgtgt 523
QY 691 ccacatgt 750
DB 524 ccacatgt 583
QY 751 attttgtttgacacacatagcagcagcagcagcagcagcagcagcagcagcagcagcagc 808
DB 584 tttttgacgt 641

RESULT 12

ID AAF28689
XX AAF28689 standard; cDNA; 1308 BP.

AC AAF28689;

XX 05-APR-2001 (first entry)

DT Human protein HP03372 coding sequence #2.

DE Human; hydrophobic domain; Immune deficiency; autoimmune disorder;
KW allergy; tissue growth; regeneration; wound healing; burn; tumour;
KW periodontal disease; thrombolytic condition; haemostatic condition;
KW infection; ss.
XX
OS Homo sapiens.
XX

PN WO200102563-A2.
XX 11-JAN-2001.
PD 16-JUN-2000; 2000WO-JP03943.
PF 02-JUL-1999; 99JP-0188835.
XX (SAGA) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
PI Kato S, Kimura T;
XX WPI; 2001-071581/08.
DR P-PSDB; AAB61608.
XX New human proteins with hydrophobic domains, useful for the treatment
PT of immune disorders, tumors, allergic conditions, thrombosis and
PT microbial infection -
XX Claim 4; Pages 122-124; 153pp; English.
XX The present invention relates to human proteins (AAB61608-AAB61617) and
CC their coding sequences (AAF28679-AAF28698). The proteins of the present
CC invention have hydrophobic domains and can be used for the treatment of
CC various immune deficiencies and disorders, such as severe combined
CC immunodeficiency (SCID), multiple sclerosis, rheumatoid arthritis,
CC autoimmune pulmonary inflammation, graft-versus-host disease and
CC Guillain-Barre syndrome. The proteins may also be useful in the treatment
CC of allergic reactions and conditions, such as asthma and in regulation of
CC haematopoiesis or lymphoid cell deficiencies. The proteins may also have
CC utility in compositions used for bone, cartilage, tendon and/or nerve
CC tissue growth or regeneration as well as wound healing and in the
CC treatment of burns. The proteins may be used in the treatment of
CC periodontal disease and in other tooth repair processes. Other uses
CC include treatment of thrombolytic and haemostatic conditions, treatment
CC or prevention of tumours and inhibiting infection by bacteria, viruses,
CC fungi and other parasites.
XX
SQ Sequence 1308 BP; 402 A; 246 C; 264 G; 396 T; 0 other;

Query Match 14.0%; Score 286; DB 22; Length 1308;
Best Local Similarity 67.4%; Pred. No. 1.5e-63;
Matches 403; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 211 gaaatgtgaattcagtcacatgccaaactaaagctgtggaccacgaagcattcatgaata 270
DB 119 gatctatgtgttatgacaaagaagaacaaatgcaaaccttgagctaaatgaata 178
QY 271 ttagtgaatccacacatcaaggctatccctgtgaggaaatgaagtcgcaactgaag 330
DB 179 ttaggcagattattcttactgggttattcttctgagagatgattgtacacaaaag 238
QY 331 atgggtatattcttctgtttaaagagattctctgagccttagtgcaacctaaagaacag 390
DB 239 atgggtatattcttctgtttaaagagattctctgagccttagtgcaacctaaagaacag 298
QY 391 gtccagcctgt 450
DB 299 ctccaaagcctgt 358
QY 451 ccaacctgcccaacatagcctgt 510
DB 359 gcaacctgcccaacacagattgt 418
QY 511 tggggacacagcaggggaacacgctgtgtctgaaacacacagacactctccatagacccaag 570
DB 419 tggggacacagcaggggaacacgctgtgtctgaaacacacacactctctaaattgtcacccgaatcac 478
QY 571 atgagttctgt 630
DB 479 cagaatactggcctcagttgt 538

QY 631 actttattttcagaaaaagccaggaagaaagatctattatgtcggctatttcacagagga 690
 Db 539 attttatcatagaaaaagcaggaagcactctactactcgttggccactcacaagga 598
 QY 691 ccacatgggtttatttcatttccaccatgcccagagctgctcagaaaaatcaaatgt 750
 Db 599 ccacatagcttttatagcattttctcaaacaccagactggtctcaaaagattagatat 658
 QY 751 attttgttagcaccatagcactgttaagcatgcaaaaaagcccgaggacaaatt 808
 Db 659 tttttgcagctccagttgtcagagtttaaacaccacaaagtccttatgaaaaaact 716

RESULT 13

AAA44349
 ID AAA44349 standard; CDNA; 617 BP.

XX AC AAA44349;

XX DT 21-AUG-2000 (first entry)

XX DE Human secreted expressed sequence tag SEQ ID NO:924.

XX KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
 KW antiulcer; osteopathic; neuroprotective; neurotropic; antipsoriatic;
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
 KW vaccine, autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
 KW tumour; infection; depression; psoriasis; ss.

XX OS Homo sapiens.

XX PN WO200021991-A1.

XX PD 20-APR-2000.

XX PF 15-OCT-1999; 99WO-US24206.

XX PR 15-OCT-1998; 98US-010436.

XX PA (GEMV) GENETICS INST INC.

XX PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 XX Merberg D, Treacy M, Bowman MR;

XX DR WPI; 2000-317938/27.

XX PT Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (SESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -

XX PS Claim 1; Page 441; 803pp; English.

XX CC AAA43426 to AAA5925 represent specifically claimed secreted expressed
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
 CC tissue sources. The SESTs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
 CC antiasthmatic; vulnary; antiulcer; osteopathic; neuroprotective;
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene
 CC therapy and in vaccines. The SESTs are useful as probes for the

CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.

XX SQ Sequence 617 BP; 181 A; 125 C; 139 G; 172 T; 0 other;

Query Match 12.6% Score 257.8; DB 21; Length 617;
 Best Local Similarity 70.7% Pred. No. 1.8e-56;
 Matches 343; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 244 ctgaggaccaggaagcattcatgaattagtgaaatcatccacatcaaggatccct 303
 Db 133 ctggaagccctgaagtgaattgaacattagtcagatgattactattgggataccaa 192
 QY 304 gtgaggatgaagtgcgaactgaagatgggtatatactcttctgtaacggattccct 363
 Db 193 atgagaatgaagtgtgactgaagatggtattattctgaagtcataagaattccct 252
 QY 364 gagccttagtgcaacctaaagaagcaggtccaggcctgtgttactcagcagatgcc 423
 Db 253 atgggaagaaaaattcgggaatacaggccagagacctgtgttttgcagatggct 312
 QY 424 tagttgaggtgtcagcaactggatttccaaacctgcccaacaaatagcctggcttcattc 483
 Db 313 tgctgcatcagccacaaactggatttccaaacctgccgaacacagccttgcttcattc 372
 QY 484 tggcagatgctgttttgacgtgtggatggggaacagcaggggaacacccctggtctcga 543
 Db 373 tggcagatgctgtgttatgtgtgtgctggcaacagcagaggaacacccctgggcaga 432
 QY 544 aacacaagacactctccatagaccagatgattcttctggccttcagttatgagatgg 603
 Db 433 gaaactgtactattaccagattcagttgaattcttggccttcagcttgatgaaatgg 492
 QY 604 ctgagtttgaccttctcgtcagtgataaaactttatttgagaaaacggcagggaaga 663
 Db 493 ctgaatatgaccttccagccacaatcgacttcattgttaagaaaactggacagaagcagc 552
 QY 664 tctattatgtcgtctattcacaggggccaccacatgggctttattgcatatttccaccatgc 723
 Db 553 tacactatgtggcattccccagggcaccacattggtttattgctctttccaccacatc 612
 QY 724 cagag 728
 Db 613 tcgag 617

RESULT 14

AAA43744

ID AAA43744 standard; CDNA; 572 BP.

XX AC AAA43744;

XX DT 21-AUG-2000 (first entry)

XX DE Mouse secreted expressed sequence tag SEQ ID NO:319.

XX KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
 KW antiulcer; osteopathic; neuroprotective; neurotropic; antipsoriatic;

[illegible]

RESULT	15	
AAF58252/C		
ID	AAF58252	standard; DNA; 936 BP.
XX		
AC	AC	
AAF58252;		
XX		
DT	24-APR-2001	(first entry)
XX		
DE	Oligonucleotide D1835.	
XX		
KW	Electron-transfer group; ETM;	mismatch; genotyping;
KW	gene expression; ss.	
XX		
OS	Synthetic.	
XX		
PN	W0200107665-A2.	
XX		
PD	01-FEB-2001.	
XX		
PF	26-JUL-2000; 2000WO-US20476.	
XX		
PR	26-JUL-1999; 99US-0145695.	
PR	17-MAR-2000; 2000US-0190259.	
XX		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.	
XX		
PI	Umek RM;	
XX		
DR	WPI; 2001-159728/16.	
XX		
PT	Nucleic acids containing electron-transfer group, u	
PT	hybridization assays, e.g. for genotyping, allowing	
PT	a single surface	
XX		
PS	Example 6; Page 127; 159pp; English.	

CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX
 SQ Sequence 936: BP; 4 A; 139 C; 10 G; 7 T; 776 other;

 Query Match 6.2%; Score 127.6; DB 22; Length 936;
 Best Local Similarity 2.7%; Pred. No. 4.7e-23;
 Matches 16; Conservative 386; Mismatches 200; Indels 0; Gaps

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 20:11:58 ; Search time 4395.77 Seconds
(without alignments)
7192.387 Million cell updates/sec

Title: US-09-333-159-45
Perfect score: 2044
Sequence: 1 gtcgaccacgcgtccgggg.....aaaaaaaaaaggcgccgcg 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_ba3:
4: gb_in1:
5: gb_in2:
6: gb_in3:
7: gb_in4:
8: gb_ov:
9: gb_pat1:
10: gb_pat2:
11: gb_ph:
12: gb_pl1:
13: gb_pl2:
14: gb_pl3:
15: gb_pl4:
16: em_ba1:
17: em_ba2:
18: em_fun:
19: em_htgo_hum:
20: em_htgo_inv:
21: em_htgo_rod:
22: em_htg_hum1:
23: em_htg_hum2:
24: em_htg_hum3:
25: em_htg_hum4:
26: em_htg_hum5:
27: em_htg_hum6:
28: em_htg_hum7:
29: em_htg_hum8:
30: em_htg_inv1:
31: em_htg_inv2:
32: em_htg_other:
33: em_htg_rod:
34: em_hum1:
35: em_hum2:
36: em_hum3:
37: em_hum4:
38: em_hum5:
39: em_hum6:
40: em_hum7:
41: em_in:
42: em_om:
43: em_or:

44: em_ov:
45: em_pat:
46: em_ph:
47: em_pl:
48: em_ro:
49: em_sts:
50: em_sy:
51: em_un:
52: em_v1:
53: gb_sts1:
54: gb_sts2:
55: gb_sts3:
56: gb_sy:
57: gb_un:
58: gb_v11:
59: gb_v12:
60: gb_htg1:
61: gb_htg2:
62: gb_htg3:
63: gb_htg4:
64: gb_htg5:
65: gb_htg6:
66: gb_htg7:
67: gb_htg8:
68: gb_htg9:
69: gb_htg10:
70: gb_htg11:
71: gb_htg12:
72: gb_htg13:
73: gb_htg14:
74: gb_htg15:
75: gb_htg16:
76: gb_htg17:
77: gb_htg18:
78: gb_htg19:
79: gb_htg20:
80: gb_htg21:
81: gb_htg22:
82: gb_htg23:
83: gb_htg24:
84: gb_htg25:
85: gb_pr1:
86: gb_pr2:
87: gb_pr3:
88: gb_pr4:
89: gb_pr5:
90: gb_pr6:
91: gb_pr7:
92: gb_pr8:
93: gb_pr9:
94: gb_rol:
95: gb_rol2:
96: gb_in4:
97: gb_pr10:
98: em_ba3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	892	43.6	181170	79	AL353113
c 2	890.4	43.6	158195	63	AC015461
c 3	630.2	30.8	76653	63	AC015506
4	527.8	25.8	2481	93	HSLYACLY
5	527.8	25.8	2481	97	HSU08464
6	527.8	25.8	2626	93	HSLAL
7	526.2	25.7	2493	97	HUMLIPCHL
8	489.2	23.9	1378	9	A26689

AL353113 Homo sapi
AC015461 Homo sapi
AC015506 Homo sapi
231690 H.sapiens
U08464 Human lysos
X76488 H.sapiens m
M74775 Human lysos
A26689 Precursor o

```

c 9 489.2 23.9 1378 9 A26690
10 470 23.0 1365 93 HSGIR
11 470 23.0 1367 9 A01046
12 470 23.0 1369 9 A12714
13 467.6 22.9 1336 9 A01157
14 467.6 22.9 1335 95 RNLIP
15 461 22.6 2358 94 MMLYACLY
16 459.8 22.5 3144 95 S81497
17 458 22.4 1417 7 BOVME
18 444.2 21.7 1137 9 AR039022
19 444.2 21.7 1137 9 AR02633
20 444.2 21.7 1140 9 AR039023
21 444.2 21.7 1140 9 AR02634
22 444.2 21.7 1146 9 AR039024
23 444.2 21.7 1146 9 AR02635
24 444.2 21.7 1198 9 A57760
25 444.2 21.7 1528 9 A57756
26 444.2 21.7 1528 9 AR039021
27 444.2 21.7 1528 9 AR02632
28 444.2 21.7 1651 7 CFLIPASE
29 441 21.6 1137 9 A39301
30 441 21.6 1140 9 A39303
31 441 21.6 1146 9 A39305
32 441 21.6 1528 9 A39300
33 388.8 19.0 1048 9 A57758
34 286 14.0 699 10 AX068249
35 286 14.0 1308 10 AX068259
36 255 12.5 148720 65 AC019353
37 255 12.5 161826 90 AL358532
38 255 12.5 172834 62 AC011277
39 151.2 7.4 76653 63 AC015506
40 140.6 6.9 401 93 HSLIPCHL04
41 140.6 6.9 1851 93 HSLIP04
42 140.6 6.9 180366 81 AL513533
43 127.6 6.2 40114 6 CEL2K6
44 122.4 6.0 137911 61 AC009769
45 122.4 6.0 204102 63 AC013826

```

ALIGNMENTS

```

RESULT 1
AL353113 181170 bp DNA HTG 08-APR-2001
LOCUS Homo sapiens chromosome 10 clone RP11-30415, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 unordered pieces.
ACCESSION AL353113.4 GI:13567970
VERSION HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 181170)
Grafham, D.
Direct Submission
Submitted (07-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:9796998.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA30415
----- Summary Statistics
Sequencing program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 180405 bases at least Q40

```

```

Consensus quality: 180692 bases at least Q30
Consensus quality: 180914 bases at least Q20
Insert size: 181070; sum-of-contigs
Insert size: 150393; agarose-fp
Quality coverage: 7.50x in Q20 bases; sum-of-contigs Quality
coverage: 9.08x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 174783: contig of 174783 bp in length
* 174784 174883: gap of 100 bp
* 174884 181170: contig of 6287 bp in length.

```

FEATURES

```

source
Location/Qualifiers
1..181170
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-30415"
/clone_lib="RPC1-11.2"
1..174783
/note="assembly_fragment:02734
clone_end:SP6
vector_side:left"
174884..181170
misc_feature
/note="assembly_fragment:02008"
BASE COUNT 57065 a 34281 c 33093 g 56630 t 101 others
ORIGIN

```

```

Query Match 43.6%; Score 892; DB 79; Length 181170;
Best Local Similarity 100.0%; Pred. No. 2.2e-213;
Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1126 agccaaactctgttaaggtacagagatgacagatgacggtccctacagcaatggagacg 1185
|||||
DB 134716 AGCCAACTCTGTAAAGGTACAGAGATGACGGTCCCTACAGCAATGGGACAG 134775
|||||
QY 1186 gaggtcaggactggcttcaaatccagaagacgtgaaatgctgctctgaggtgacca 1245
|||||
DB 134776 GAGGTGAGGACTGGCTTCAATCCAGAACGCTGAAATGCTGCTCTGAGGTGACCA 134835
|||||
QY 1246 accctatcaccaagaataattcttcgaatgggctacggtgatttcattctggtgttgg 1305
|||||
DB 134836 ACCTCATCTACCAATAAGAAATATTCCTGAATGGGCTCACGGGATTCATCTGGGGTTGG 134895
|||||
QY 1306 atgctctcacccatgtacaaatgaaatccatctcctgagcagagagagagaccacc 1365
|||||
DB 134896 ATGCTCTCACCGTATGTACAATGAATCATCTCATCTGATGACGAGAGAGACCAACC 134955
|||||
QY 1366 ttccccaggagacggtgagggccgtattgtgaagcatctgacactgacgtattcttaggaca 1425
|||||
DB 134956 TTTCCTCAGGAGCGGTGTAGGGCGGTATTGTGAAGCATCTGCACACTGACGATCTTAGGACA 135015
|||||
QY 1426 accctctgagggatggggttaggaccatgaagcagaataaccgagagagagagaccatag 1485
|||||
DB 135016 ACCTCTCAGGGATGGGGCTTAGGACCCCATGAAGGCAAAATTCGGAGAGAGAGACCTAG 135075
|||||
QY 1486 tatacatcttttcagattccctgcaactggcactaaatccgacactacattacattttt 1545
|||||
DB 135076 TATCATATTTTCAGATTCCTGCGACTGGCACTAAATCCGACACTTACATTATTTT 135135
|||||
QY 1546 ttctgtaaattaaagtactatttaggttaaatagaggtttttagctattatatctt 1605
|||||
DB 135136 TTCTCTGTAATAAAGTACTATTATAGGTAATAAGAGGTTTGTGTATGCTATTATATATCT 135195
|||||
QY 1606 accatcttgagggttaggttttaccctgatgaccagagaaaatctatagacattctctatc 1665
|||||

```

```
Db 135196 ACCATCTTTGAAGGGTAGCTTTTACCTGATAGCCAGAAAATATCTAGACATCTCTATATC 135255
QY 1666 attcaggtaaatctctttaaacacacctattgttttttttataaagccatatttttggagca 1725
|||||
Db 135256 ATTCAAGTAAATCTCTTTAAACACACCTATTCTTTTCTATAGCCATATTTTGGAGCA 135315
|||||
QY 1726 ctcaagataaatggcaattgggacagatatattgaggtctgagctgtggtgattttgttg 1785
|||||
Db 135316 CTAAAGTAAATGGCAATTTGGACAGATATTGAGGTCTGGAGTCTGGGATTTATGTTG 135375
|||||
QY 1786 accttgacaaaataagctagacattttcaccttctgttgcacagagacataacactacctc 1845
|||||
Db 135376 ACTTTGACAAAATAAGCTAGACATTTTTCACCTTGTGTCACAGACAGACATAACACTACCTC 135435
|||||
QY 1846 aggaagctgagctgttttaagagacaaacaaacaaacaaacaaacaaacaaacaaacaaac 1905
|||||
Db 135436 AGGAAGCTGAGCTGCTTTTAAAGACAAACAAACAAACAAACAAACAAACAAACAAACAA 135495
|||||
QY 1906 tctatttaagcattctcagaataagccaaagtatttatttatttatttatttatttatttattt 1965
|||||
Db 135496 TCTATGTTAAGCATCTCTCAGATATAGGCCAAGTTTATGTTGATCTCAGGGAAGAAAA 135555
|||||
QY 1966 tttatagatgtttatgatttctccaaataaatgattcttcattctacataaa 2017
|||||
Db 135556 TTTTATAGATGTTTATGATGTTCTCCATAAATGATTCGATTCGATTCATACATAAA 135607
|||||

RESULT 2
AC015461/c
LOCUS
DEFINITION Homo sapiens clone RP11-115N17, WORKING DRAFT SEQUENCE, 10
ACCESSION AC015461
VERSION AC015461.3 GI:9502445
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-115N17
Unpublished
2 (bases 1 to 158195)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgaiter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczký,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McDonald,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 27, 2000 this sequence version replaced gi:9112670.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L1907
Center clone name: 115_N_17
```

```
----- Summary Statistics
Sequencing vector: M13; M77615; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152466 bases at least Q40
Consensus quality: 154897 bases at least Q30
Consensus quality: 156006 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 157295; sum-of-contigs
Quality coverage: 5.8 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1002: contig of 1002 bp in length
* 1003 1102: gap of 100 bp
* 1103 2388: contig of 1286 bp in length
* 2389 2488: gap of 100 bp
* 2489 3525: contig of 1037 bp in length
* 3526 3625: gap of 100 bp
* 3626 16527: contig of 12902 bp in length
* 16528 16627: gap of 100 bp
* 16628 28636: contig of 12009 bp in length
* 28637 28736: gap of 100 bp
* 28737 46598: contig of 17862 bp in length
* 46599 46698: gap of 100 bp
* 46699 71912: contig of 25214 bp in length
* 71913 72012: gap of 100 bp
* 72013 96612: contig of 24600 bp in length
* 96613 96712: gap of 100 bp
* 96713 128331: contig of 31619 bp in length
* 128332 128431: gap of 100 bp
* 128432 158195: contig of 29764 bp in length.
Location/Qualifiers
1. 158195
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-115N17"
/clone_lib="RPC1-11 Human Male BAC"
1. 1002
/note="assembly_fragment"
1103 2388
/note="assembly_fragment"
2489 3525
/note="assembly_fragment"
3626 16527
/note="assembly_fragment"
16628 28636
/note="assembly_fragment
clone_end:r7
vector_side:right"
28737 46598
/note="assembly_fragment"
46699 71912
/note="assembly_fragment"
72013 96612
/note="assembly_fragment"
96713 128331
/note="assembly_fragment"
128432 158195
/note="assembly_fragment
clone_end:SP6
vector_side:left"
905 others
```

BASE COUNT 48628 a 30917 c 30188 g 47557 t 905 others
ORIGIN

Query Match 43.68; Score 890.4; DB 63; Length 158195;
 Best Local Similarity 99.98; Pred. No. 5.4e-213;
 Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1126	agcaactcctgtaaggtacagagtcagagatagcggtccctacagcaatgtggagac	1185
Db	19482	AGCCAACTCCGTAAAGTCAGAGTCAGATAGCGGTCCCTACAGCAATGTGGACAG	19423
Qy	1186	gagtcaggactgcttccaatcagagacgctgaaatgctgctctgagtgacca	1245
Db	19422	GAGTCAAGACTGCTTCAATCAGAGACGCTGAAATGCTGCTCTGAGTGACCA	19363
Qy	1246	acctcattaccataagaataatctgaatgggctcagtggtatttcattcggtttgg	1305
Db	19362	ACCTCACTACCAATAAGAATAATCTGAATGGGCTCAGTGGATTTTCATCTGGGGTTGG	19303
Qy	1306	atgctctccacgtatgtacaaataatccatccatctgtagcagcagagagacacac	1365
Db	19302	ATGCTCTCCACGCTATGTACAAATCAATCATCTCATCTGAGCAGTCTTAGGACA	19243
Qy	1366	ttcccaggagcgtgtgagccgtattgtgaagcattctgacactgacgtatcttaggaca	1425
Db	19242	TTTCCAGGGACGCTGTGAGGCCCTATTGTGAGCATCTGACACTGACGATCTTAGGACA	19183
Qy	1426	acctcctgagggatggggctaggaccatgaagcagaataattacgagagagacctag	1485
Db	19182	ACCTCCTGAGGGATGGGGCTAGGACCCATGAAGCAGCAATATACAGAGAGACCTAG	19123
Qy	1486	tatacttttccagatccctgcacttggcactaaatccgacacttaccatttcatcttt	1545
Db	19122	TATACATTTTTCAGATCCCTGCACCTTGGCATAATCCGACACTTACATTACATTTT	19063
Qy	1546	ttctgtaaataaagctactattaggtaaatagaggtttgtgattgcattatatattct	1605
Db	19062	TTTCTGTAATTAAGTACTATTAGTAAATAGAGGTTTGTATGCTATTATATATCT	19003
Qy	1606	acctcttgagggtggttttacctgtagccagaaatctctagcattctcttatct	1665
Db	19002	ACCATCTGAAGGTAGGTTTACCTGATAGCCAGAAATATCTAGACATCTCTATATC	18943
Qy	1666	attcaggtaaatctcttaaaacacctattgttttctataagccattttttggagca	1725
Db	18942	ATTGAGTAAATCTCTTTAAACACCTATTGTTTTTCTATAAGCCATATTTTGGAGCA	18883
Qy	1726	ctaaagtaaaatggcaatgggacagatattgaggtctgaggtctggtattgtttg	1785
Db	18882	CTAAAGTAAATGGCAATTTGGGACAGATATTGAGGCTCTGGAGTCTGTGATTATTGTG	18823
Qy	1786	acttgacaaaataagctagacatttccactgttgcacagagacataacactacctc	1845
Db	18822	ACTTGCACAAATTAAGTAGACATTTTACCTGTTGCTCCACAGAGACATAACACTACCTC	18763
Qy	1846	aggaagctgagctgctttaaaggacaaacaaacaaatcagtggttacagtatggtgaaa	1905
Db	18762	AGGAAGCTGAGCTGCTTTAAGGACAAACAAACAAATCAGTGTACAGTATGATGAAA	18703
Qy	1906	tctatgtaagcattctcagaataaagccaaagttttatagttgcattctcgggaagaaa	1965
Db	18702	TCTATGTAAACATCTCTCAGAATAAGCCCAAGTTTATAGTTGCTATCTCAGGGAAGAAA	18643
Qy	1966	ttttataggtgtttatgagcttcccaataaataatgcattctgcattacataaa	2017
Db	18642	TTTTATAGGATGTTTATGAGTCTTCCAAATAATGCATTCGATTACATAAA	18591

RESULT 3
 LOCUS AC015506/c
 DEFINITION Homo sapiens clone RP11-21124, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC015506
 VERSION AC015506.3 GI:9124007
 KEYWORDS HTG; HTGS_PHASED.
 SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 76653)

AUTHORS

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Unpublished

REFERENCE

2 (bases 1 to 76653)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collins, A.,
 Cooke, P., DeArillano, K., Dewar, K., Domino, M., Donnelly, L., Doyle, M.,
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gardyna, S., Grant, G., Hages, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tasfaye, S., Tirell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, J. M.

TITLE

Direct Submission

JOURNAL

Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6468098.
 All repeats were identified using RepeatMasker:

COMMENT

Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L4028

Center clone name: 21_1_24

Center clone name: 21_1_24

* NOTE: This record contains 87 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

* 795: contig of 795 bp in length

* 796 895: gap of 100 bp

* 896 1656: contig of 761 bp in length

* 1657 1756: gap of 100 bp

* 1757 2558: contig of 802 bp in length

* 2559 2658: gap of 100 bp

* 2659 3434: contig of 776 bp in length

* 3435 3534: gap of 100 bp

* 3535 4301: contig of 767 bp in length

* 4302 4401: gap of 100 bp

* 4402 5199: contig of 798 bp in length

* 5200 5299: gap of 100 bp

* 5300 6085: contig of 786 bp in length

* 6086 6185: gap of 100 bp

* 6186 6972: contig of 787 bp in length

* 6973 7072: gap of 100 bp

* 7073 7844: contig of 772 bp in length

* 7845 7944: gap of 100 bp

* 7945 8731: contig of 787 bp in length

* 8732 8831: gap of 100 bp

* 8832 9625: contig of 794 bp in length

* 9626 9725: gap of 100 bp

* 9726 10519: contig of 794 bp in length

* 10520 10619: gap of 100 bp

```

* 10620 11417: contig of 798 bp in length
* 11418 11517: gap of 100 bp
* 11518 12296: contig of 779 bp in length
* 12297 12396: gap of 100 bp
* 12397 13180: contig of 784 bp in length
* 13181 13280: gap of 100 bp
* 13281 14065: contig of 785 bp in length
* 14066 14165: gap of 100 bp
* 14166 14949: contig of 784 bp in length
* 14950 15049: gap of 100 bp
* 15050 15822: contig of 773 bp in length
* 15823 15922: gap of 100 bp
* 15923 16718: contig of 796 bp in length
* 16719 16818: gap of 100 bp
* 16819 17621: contig of 803 bp in length
* 17622 17721: gap of 100 bp
* 17722 18517: contig of 796 bp in length
* 18518 18617: gap of 100 bp
* 18618 19390: contig of 773 bp in length
* 19391 19490: gap of 100 bp
* 19491 20253: contig of 763 bp in length
* 20254 20353: gap of 100 bp
* 20354 21134: contig of 781 bp in length
* 21135 21234: gap of 100 bp
* 21235 22028: contig of 794 bp in length
* 22029 22138: gap of 100 bp
* 22139 22905: contig of 777 bp in length
* 22906 23005: gap of 100 bp
* 23006 23787: contig of 782 bp in length
* 23788 23887: gap of 100 bp
* 23888 24648: contig of 761 bp in length
* 24649 24748: gap of 100 bp
* 24749 25529: contig of 781 bp in length
* 25530 25629: gap of 100 bp
* 25630 26414: contig of 785 bp in length
* 26415 26514: gap of 100 bp
* 26515 27308: contig of 794 bp in length
* 27309 27408: gap of 100 bp
* 27409 28165: contig of 761 bp in length
* 28170 28269: gap of 100 bp
* 28270 29059: contig of 790 bp in length
* 29060 29159: gap of 100 bp
* 29160 29953: contig of 794 bp in length
* 29954 30053: gap of 100 bp
* 30054 30834: contig of 781 bp in length
* 30835 30934: gap of 100 bp
* 30935 31716: contig of 782 bp in length
* 31717 31816: gap of 100 bp
* 31817 32590: contig of 774 bp in length
* 32591 32690: gap of 100 bp
* 32691 33471: contig of 781 bp in length
* 33472 33571: gap of 100 bp
* 33572 34334: contig of 763 bp in length
* 34335 34434: gap of 100 bp
* 34435 35215: contig of 781 bp in length
* 35216 35315: gap of 100 bp
* 35316 36083: contig of 768 bp in length
* 36084 36183: gap of 100 bp
* 36184 36962: contig of 779 bp in length
* 36963 37062: gap of 100 bp
* 37063 37806: contig of 744 bp in length
* 37807 37906: gap of 100 bp
* 37907 38682: contig of 776 bp in length
* 38683 38782: gap of 100 bp
* 38783 39570: contig of 788 bp in length
* 39571 39670: gap of 100 bp
* 39671 40463: contig of 793 bp in length
* 40464 40563: gap of 100 bp
* 40564 41353: contig of 790 bp in length
* 41354 41453: gap of 100 bp
* 41454 42216: contig of 763 bp in length
* 42217 42316: gap of 100 bp
* 42317 43093: contig of 777 bp in length

```

```

* 43094 43193: gap of 100 bp
* 43194 43966: contig of 773 bp in length
* 43967 44066: gap of 100 bp
* 44067 44841: contig of 775 bp in length
* 44842 44941: gap of 100 bp
* 44942 45707: contig of 766 bp in length
* 45708 45807: gap of 100 bp
* 45808 46598: contig of 791 bp in length
* 46599 46698: gap of 100 bp
* 46699 47457: contig of 759 bp in length
* 47458 47557: gap of 100 bp
* 47558 48363: contig of 806 bp in length
* 48364 48463: gap of 100 bp
* 48464 49263: contig of 800 bp in length
* 49264 49363: gap of 100 bp
* 49364 50154: contig of 791 bp in length
* 50155 50254: gap of 100 bp
* 50255 51047: contig of 793 bp in length
* 51048 51147: gap of 100 bp
* 51148 51939: contig of 792 bp in length
* 51940 52039: gap of 100 bp
* 52040 52823: contig of 784 bp in length
* 52824 52923: gap of 100 bp
* 52924 53714: contig of 791 bp in length
* 53715 53814: gap of 100 bp
* 53815 54604: contig of 790 bp in length
* 54605 54704: gap of 100 bp
* 54705 55491: contig of 787 bp in length
* 55492 55591: gap of 100 bp
* 55592 56385: contig of 794 bp in length
* 56386 56485: gap of 100 bp
* 56486 57264: contig of 779 bp in length
* 57265 57364: gap of 100 bp
* 57365 58144: contig of 780 bp in length
* 58145 58244: gap of 100 bp
* 58245 59001: contig of 757 bp in length
* 59002 59101: gap of 100 bp
* 59102 59875: contig of 774 bp in length
* 59876 59975: gap of 100 bp
* 59976 60753: contig of 778 bp in length
* 60754 60853: gap of 100 bp
* 60854 61636: contig of 783 bp in length
* 61637 61736: gap of 100 bp
* 61737 62508: contig of 772 bp in length
* 62509 62608: gap of 100 bp
* 62609 63382: contig of 774 bp in length
* 63383 63482: gap of 100 bp
* 63483 64260: contig of 778 bp in length

```

Query Match 30.88; Score 630.2; DB 63; Length 76653;
 Best Local Similarity 99.48; Pred. No. 1.4e-147;
 Matches 632; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Qy 1126 agccaaacctgtaaggtacagagatgacaggtccctacagcaatgtggacag 1185
Db 30724 AGCCAACTCTGTAAAGGTACAGAGTCAGAGATATGACGGTCTTACACACATGTGGACAG 30665
Qy 1186 gagtcaggactgcttcaaatccagaagacgtgaaatgctgtctctgagtgacca 1245
Db 30664 GAGTCAGGACTGGCTTTCAAATCCAGAAGACGTGAAATGCTGCTCTCTGAGGTGACCA 30605
Qy 1246 acctcatctaccataagaattctgaatgggtcactgatttcctgaggtttgg 1305
Db 30604 ACCTCATCTACCAATGAAGATATTCCTGAATGGGTCTGAGTTCATCTGAGGTGGTGG 30545
Qy 1306 atgctctaccgtatgtacaatgaatcatccatctgatcagcagagagagacaaacc 1365
Db 30544 ATGCTCTCACCCTATGTATACATCAATCAATCATCCATCTGATGCCAGGAGGAGACCAACC 30485
Qy 1366 ttcccaggagcgtgtgagccgtatttgtgaagcatcttgacactgacatcttaggaca 1425
Db 30484 TTTCCAGGAGCGTGTGAGGCCCGTATTGTGAAGCATCTGACACTGACATCTTAGGACA 30425

```


Qy 638 tttcagaaaaacggccagagaaagatctctattatgtcggtattcaccagggccaccacat 697
Db 517 TCTGAATAAAATGCGCAAGAACAAAGGTATATATGGTTCATCTCTCAAGCCACCATAT 576
Qy 698 gggctttattgcatcttccaccatgcccagagctggctcagaaaaatcaaaatgtatttgc 757
Db 577 AGGTTTATACATTTTCACAGATCCCTGAGCTGGCTTAAAGGATTAAATGTTTTTTCG 636
Qy 758 tttagcaccatagaccactgttaagcatgcgcaaaagcccccgggagccaaaatttttggct 817
Db 637 CTTGGGCTCTGCTGCTCCGCTGCTTCTGTACTAGCCCTATGCCCCAAATTAGGACCAT 696
Qy 818 gccagatatgatcatcaaggattgtttggcaaaaaagaatttctgtatcagaccagatt 877
Db 697 ACCAGATCATCTCATTAAGACATTATTGGAGACAAAGAATTTCTTCCCCAGAGTGGCT 756
Qy 878 tctcagacaactgttattacccttggccaggtgattcttgatcagatttggtagtaa 937
Db 757 TTTGAAGTGGTGGGTACCCAGCTTTGCACTCATGTCTACTGAAGGAGCTCTGTGGAAA 816
Qy 938 tatcatgttacttcttgggtgattcaacacccaacaaatgatgaacatgagccgagcaagtgt 997
Db 817 TCTCTGTTTCTTCTGTGTGGATTATAGAGAAATTTAAATATGTCTAGAGTGGATGT 876
Qy 998 atatgctgcccacactcttctggaacatctgtgcaaaaattattacactggagccagggc 1057
Db 877 ATATACAAACACATTTCTCTGCTGGAACTTCTGTSCAAACATGTTACACTGGAGCCAGG 936
Qy 1058 agtgaattctgtgaactcgggcatcttgactggggagtgagaccacaaaatctggaaaa 1117
Db 937 TGTAAATTCCAAAGTTTCAAGCTTTGACTGGGGAGGAGCTGCCAAGAATATTATTCA 996
Qy 1118 atgcaatcagccaactcgttaaggtacagagtacagatatgacggtccctcagcaat 1177
Db 997 TTACAACAGAGTTATCTCCACATACAAATGTGAAGGACATGCTTGTGCCGAGTGCAGT 1056
Qy 1178 gtgacagggaggtcagagactggtttcaaatccagaagagcgtgaaatgtgctctctga 1237
Db 1057 CTGGAGGGGGGTCCAGCTGGCTTGCAGATGTCTACGACGTCAATATCTTACTGACTCA 1116
Qy 1238 ggtgaccaacctcctaccataagaattcctggaatggctcagctgatttcatctg 1297
Db 1117 GATCACCAACTTGTGTCTCATGAGACATTCGGAAATGGAGCATCTTCACTTCATTG 1176
Qy 1298 ggggttgatgctcaccgcatgtacaaatgaaatcatccatctgtatgagcaggagga 1357
Db 1177 GGGCTGGATGCCCTTGGAGGCTTATATAAAATTTATTAATCTAATGAGGAATATCA 1236
Qy 1358 gaccacact 1366
Db 1237 GTGAAGCT 1245

RESULT 8
A26689
LOCUS A26689 1378 bp DNA PAT 05-APR-1995
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE Benicourt,C., Blanchard,C. and Junien,J.L.
AUTHORS Recombinant gastric lipase from rabbit and pharmaceutical
TITLE compositions
JOURNAL Patent: EP 0542629-A 9 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES Location/Qualifiers
source 1. .1378
/organism="synthetic construct"

BASE COUNT 397 a 287 c 276 g 418 t
ORIGIN
Query Match 23.9%; Score 489.2; DB 9; Length 1378;
Best Local Similarity 65.0%; Pred. No. 2.5e-112;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
Qy 250 acccagaagcattcatgaatattagtaaaatcccaacatcaagcattccctctgtaagg 309
Db 136 ACCCTGAAGTGAAATATATATAGTCAGATGATTTCTACTGGGATACCAAGTGA 195
Qy 310 aatatgaagtcgcaactgaagatgggtatatctctttgtttaacaggaattcctcgagcc 369
Db 196 AATATGAAGTTGTGACTGAAGATGGTTATATCTTGAAGTTAACAGAATTCCTTATGGGA 255
Qy 370 tagtgcaacctagaagacagggttccaggcctgtgggtgttactcagcatggcctagttg 429
Db 256 AGAAAAATTTCAGGGAACAGAGCCAGAGACCGCTCGTATTATTTTCAGCATGGTTGCTTG 315
Qy 430 gaggtgctagcaactgatttccaaacctgcccaacatagcctgggttcttattctgcag 489
Db 316 CATCAGCCTCAAACTGATCTCCAACTGCCAAACACAGCCTGGCCTTTATTCTGGCAG 375
Qy 490 atgctgggttttgagctgtggtgaggggaacagcagggggaacgcctgctcgaacacaca 549
Db 376 ATGCTGGTTATGTGTGTGGTGGGAACAGAGAGAAATACCTGTCCAGAGAANAAT 435
Qy 550 agacactctccatagaccagaagatgagttctgggcttccagttatgatgagatggctaggt 609
Db 436 TATACTATTACCAGACTCCGTTGAATTCGGGCTTTTAGCTTTTGATGAATGGCTAAAT 495
Qy 610 ttgcctctctgctgagtaaaactttatttgagaaaacggcggcagaaaagatctatt 669
Db 496 ATGACCTTCCAGCCACAAATTCAGTTCAATGTAAGGAAACTGGACAGAAAGCTGCAC 555
Qy 670 atgtcggtctattcacagggccaccactgggctttattgcattttccaccatgcccagc 729
Db 556 ATGTTGTGTCATCTCCAGGGCACCACCATTGTTTATTGCTTTTCTACAAATCCCAAGC 615
Qy 730 tggctcagaaaaatcaaaatgtatttcttagcaccatagccactgttaagcatgcaa 789
Db 616 TGCTGAAAGAAATCAAAACCTTTTATGCATTAGCTCCAGTCGCCACTGTGAAGTATCAA 675
Qy 790 aaagcccccgggacaaaatgttctgctgcagatatgatcaagggatgttttgca 849
Db 676 AAGCCCTTGTAAACAAACTTAGGTTTATCTCTCCAACCAATGTTCAAGATTATATTTGGTG 735
Qy 850 aaaaaaattctgtatcagaccagatttctcagacaaactgtttatttaccctttgtggcc 909
Db 736 ACAAAATATTCTACCCACACAAATTTCTTGATCAATTTCTGCCACTCAAGTGTGTTCCC 795
Qy 910 aggtgattctgtacagattgttagtaaatatcatctgttacttctctgggtggatcaaccca 969
Db 796 GTGAGACACTGAAATGTAATTTGCAGCAATGCCCTATTATATCATTTGTGGCTTTGCACAGC 855
Qy 970 acaatatgaacatgagccgagcaagtgtatatgtctgccacactcttctgtggaacatctg 1029
Db 856 CAAAC-TTGAACATGAGTCGCTTGGATGTGTACGTGTGCACATAATCCGGCAGGAACACTCAG 915
Qy 1030 tgcataatattctacatgagccagggcagtgatgtattctgtgnaactccgggcatattgact 1089
Db 916 TTCAAAACATGCTGCACACTGGACCCAGGCTGTTAAATCTGGAATTTTCAAGCTTTTAAT 975
Qy 1090 ggggagtgagacaaaatctggaataatgcaatcagcaactcctgtgaaggtacagag 1149
Db 976 GGGG-AGTCCAGCTCAGAATGTAGTGCACATTCAAATAGCCACACACTCCCTACTACAAATG 1035
Qy 1150 tcagagatatgacgggtccctcacagcaatgtggacagaggtcagaggtggcttttcaaatc 1209
Db 1036 TGACCCCATGATGTGCCAATTTGCACTGTGGAGTGGTGGCAATGACTGGTTGCTGCTGACC 1095

/protein_id="CAA29413.1"
 /db_xref="GI:758063"
 /db_xref="SWISS-PROT:P07098"
 /translation="MWLLTMSLISVLGTHGLFGKLPSPVETMNTISOMITYWGY
 PNEYEVYEDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLS
 AFLDAGYDVLGNSGRNTWARRNLIYSPDSVEFWAFSDEMAKYDLDPATIDFLVK
 TGOKOLHYVGHSGOTTIGFIAFSTNPSLAKRIKTFYALAPVATVYTKSLINKLRFVP
 QSLFKIFGDKIFYPHNFDDQLATEVCSREMLNLLCSNALFIICGFSKFNFTSRLD
 VYLSPAGTQVONPHFTQAVKSGKFOAYDMGSPVQNMHYDQSPPIYNTAMNRP
 IAWNGKDLADPDQVGLLPLKPLNLIYHKEIPYNHLDFIAMDAPQEVYNDIVSM
 ISDDKK"
 63..1241

CDS

/codon_start=1
 /product="gastric lipase precursor"
 /protein_id="CAA29414.1"
 /db_xref="GI:758064"
 /db_xref="SWISS-PROT:P07098"
 /translation="MASLISVLGTHGLFGKLPSPVETMNTISOMITYWGYNBEYE
 VYTEDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLSAFLIAD
 AGDYVWIGNSGRNTWARRNLIYSPDSVEFWAFSDEMAKYDLDPATIDFLVKTGOKOL
 HYVGHSGOTTIGFIAFSTNPSLAKRIKTFYALAPVATVYTKSLINKLRFVPSLFKE
 IFGDKIFYPHNFDDQLATEVCSREMLNLLCSNALFIICGFSKFNFTSRLDYLSHN
 PAGTSVQNMHYDQAVKSGKFOAYDMGSPVQNMHYDQSPPIYNTAMNRP
 GKDLADPDQVGLLPLKPLNLIYHKEIPYNHLDFIAMDAPQEVYNDIVSMISDDKK"

sig_peptide 63..101
 /product="pot.signal peptide (AA 1-13 to -1)"
 mat_peptide 102..1238
 /product="gastric lipase (AA 1-379)"
 BASE COUNT 396 a 284 c 268 g 417 t
 ORIGIN

Query Match 23.0%; Score 470; DB 93; Length 1365;
 Best Local Similarity 64.2%; Pred. No. 1.7e-107;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 244 ctgtggccagagaagcattcatgaatttagtgaaatcatccacatcaaggctatccct 303
 DB 121 CTGGAAGCCCTGAAGTGTGACTGAAGATGTTATCTTGAAGTCAATGAAATCCCT 240
 QY 304 gtgggaatgaagtcgaactgaagtggtatctcttctgtttaaagcagattcctc 363
 DB 181 ATGAAGATATGAAGTGTGACTGAAGATGTTATCTTGAAGTCAATGAAATCCCT 240
 QY 364 gagcctagtgaacctaaagaagcaggttccagggctgtggttactcagcagtgcc 423
 DB 241 ATGGGAAGAAAAATTCAGGGAATACAGCCAGAGACCTGTGTGTTTTCACAGATGGT 300
 QY 424 tagttgagtgtagcaactgagttcccaacctgcccacaataagctgggtcttcattc 483
 DB 301 TGCTTGATCAGCCACACAACTGGATTTCCACCTGCCGAACACAGCCTTGCTTCATTC 360
 QY 484 tggcagatcgtgtttgacgtgtgagtgagggaacagcaggggaacgcctgtctcgaa 543
 DB 361 TGGCAGATGCTGGTATGATGTGTGGTGGCAACAGCAGAGAAACACCTGGGCCAGAA 420
 QY 544 aacacagacactctccatcacaagaatgagttcttgggtttcagttatgatgagatgg 603
 DB 421 GAAACTTGATATTCACAGATTCAGTTGATTCAGCTTTCAGCTTTCAGTGAATGG 480
 QY 604 ctaggtttgacctctctgcagtgataaactttatcttcgagaaaacggccaggaagaaga 663
 DB 481 CTAAATATGACCTTCCAGCCACATCGACTTCATTTGTAAGAAAAAAGTGGACAGCAGC 540
 QY 664 tctattatgcgctattcagggcacccatggcttattgatttccaccatgc 723
 DB 541 TACACTATGTGGCCATCTCCAGGGCACCACTTGGTTTATGTGCTTTTCCACCAATC 600
 QY 724 cagagctggctcagaaaaatcaaatgtattttgcttagcaccatagcactgttaagc 783
 DB 601 CCAGCTGGCTAAAGAAATCAAACTCTATGCTTAGCTCTCTGTGGCCACTGTGAAGT 660

QY 784 atgcaaaaagccccgggacaaaatttttggctgcccagatatgatgatcaagggtattgt 843
 DB 661 ATACAAAAGCCCTATATAACAAACAACTAGATTCTTCTCAATCCCTCTCAAGTTATAT 720
 QY 844 ttggcaaaaagaattctctgatacagacagatttctcagacacactgttattacattt 903
 DB 721 TTGGTACAAAATATTCTTCCACACAACTTCCTTGGATCAATTTCTTGGTCTGAAGTGT 780
 QY 904 gtggcagggtgattcttctgatacagattgttagtaatatcatgttactctgggtgattca 963
 DB 781 GCTCCCGTGAGATGCTGAATCTCTTGGAGCAATGCCCTATTATTAATTTGGATTG 840
 QY 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtgtccacacactcttgcggaa 1023
 DB 841 ACAGTAAGAACTTTAAACACAGTTCGCTGGATGTGTATCTATCATATAAATCCACAGCAA 900
 QY 1024 catctgtgcaaaattctcactgagccagcagcagtgattcttctggtgaactccgggcat 1083
 DB 901 CTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTAAAGTCTGGGAAATTCACAGCTT 960
 QY 1084 ttgactgggggagtgagacaaaattctgaaaaatgcaatcagccaaactcctgtaaggt 1143
 DB 961 ATGACTGGGGAAGCCCACTGATCAGATAGGATGACTATGATGATCCCACTCCCTACT 1020
 QY 1144 acagagtcagagatatgacggtccctacagcaatgtggcagaggtcaggaactggcttt 1203
 DB 1021 ACANTGTGACAGCATGAATGTACCAATTCAGTGTGGAACGGTGGCAAGGACCTGTGG 1080
 QY 1204 caaatccagaaagcgtgaaaatctctctctgagtgagcgaacccatcattaccataaga 1263
 DB 1081 CTGACCCCAAGATGTGTGGCTTTTCTCCAAACTCCCAATCTATTATTACCAAGG 1140
 QY 1264 atattctgaatgggtcactggtgatttctctgggtttgggtctccctccacgtatgt 1323
 DB 1141 AGATTCTTTTACATCACTTGGACTTTATCTGGCAATGGATGCCCTTCAGAACTTT 1200
 QY 1324 acaatgaatcatccactctgat 1345
 DB 1201 ACAATGACATTTGTTCTATGAT 1222

RESULT 11

LOCUS AC1046 1367 bp mRNA PAT 16-MAR-1993
 DEFINITION H.sapiens mRNA for human gastric lipase.
 ACCESSION AC1046
 VERSION AC1046.1 GI:344241
 KEYWORDS gastric lipase.
 SOURCE human.
 ORGANISM Hsomo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1367)
 AUTHORS POLYPEPTIDE AND POLYPEPTIDE COMPOSITION
 TITLE Patent: WO 8601532-A 5 13-MAR-1986;
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..1367
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 sig_peptide 47..103
 CDS 47..1243
 /note="gastric lipase"
 /product="lipase"
 /protein_id="CAA00125.1"
 /db_xref="GI:344242"
 /db_xref="SWISS-PROT:P07098"
 /translation="MWLLTMSLISVLGTHGLFGKLPSPVETMNTISOMITYWGY
 PNEYEVYEDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLS
 AFLDAGYDVLGNSGRNTWARRNLIYSPDSVEFWAFSDEMAKYDLDPATIDFLVK
 TGOKOLHYVGHSGOTTIGFIAFSTNPSLAKRIKTFYALAPVATVYTKSLINKLRFVP
 QSLFKIFGDKIFYPHNFDDQLATEVCSREMLNLLCSNALFIICGFSKFNFTSRLD

Db 303 TGCTGATGACGACAAACTGGATTCCAACTGCGGAAACACAGCTTGCCCTTCAATC 362
 QY 484 tggcagatgctgtgtttgagctgtgagtgagggaacagcagggaacagcctgtgtctcgaa 543
 Db 363 TGGCAGATGCTGTGTATGATGTGTGCTGGGCAACAGCAGGAAACACCTTGGGCCAGAA 422
 QY 544 aacacaaagacactctccatagacacagatgagttctgggtttcagttatgagatgag 603
 Db 423 GAAACTTGTACTATTACACAGATTCAGTTGAATTCCTGGCTTTCAGCTTTGATGAAATGG 482
 QY 604 ctgagttgacctctctgagtgataaaactttatttgcagaaaacagcagggaagaaaga 663
 Db 483 CTAATATGACCTTCCAGCCACAACTGACATTCATTGTAAGAAACCTGGACAGACGAGC 542
 QY 664 tctattatgctggctattcacaggggaccaccactgggtttattgcattttccaccatgc 723
 Db 543 TACACTATGTGGCCATTCCAGGCGACACCACTTGGTTTATGCTTTTCCACCAATC 602
 QY 724 cagagctggctcgaaatacaaatgtattttcttttagcaccatagccactgttaagc 783
 Db 603 CCAGCCTGGCTAAAGAAATCAAAACCTCTCTATGCTCTAGCTCTGTTGCCACTGTGAAGT 662
 QY 784 atcacaagaagcccgagcaaatattttgtgctgcagatatgatcaagggattgt 843
 Db 663 ATACAAAAGCCCTATTAACAACTTAGATTGTGCTCAATCCCTCTTCAAGTTTATAT 722
 QY 844 ttggcaaaaaaagaattctgtatcacagcagatttctcacagcaactgtttatttaccctt 903
 Db 723 TTGCTGACAAATATCTACCCACACAACTTCTTTGATCAATTTCTGTACTGAAGTGT 782
 QY 904 gtggccagggtatctctgacagattgttagtaataatcatgttactctggtggattca 963
 Db 783 GCTCCCTGAGATGCTGAATCTCTTGTGACGAATGCTTATTAATAATTTGTGGATTG 842
 QY 964 acaccaacaatatgaacatagccgagcaagtgtatatgtgcccacactcttgcaggaa 1023
 Db 843 ACAGTAAGAACTTTAACACAGTGGCTTGGATGTGTATCATACATAATTCACAGCAGGAA 902
 QY 1024 catctgtgcaaaatattctacactgagccagcagtgagttctgtggaactccgggcat 1083
 Db 903 CTCTGTTCAAACATGTTCCATTGNCACCGAGCTGTTAGTCTGGGAAATTCGAAGCTT 962
 QY 1084 ttgactgggggagtgagacaaatactggaaaaatgcaatcagccaaactctgtgaagt 1143
 Db 963 ATGACTGGGGAAGCCAGTTTCAGAAATAGGATGCATATGATCAGTCCCAACCTCCCTACT 1022
 QY 1144 acagagtcagagatatgacgtccctacagcaatgtggacagaggtcaggaactggcttt 1203
 Db 1023 AATGTGACAGCCATGATGTACCAATTCAGATGTGGAAACGGTGGCAAGGACCTGTTGG 1082
 QY 1204 caaatcagagacgctgaaatgctgctctgaggtgacccaaactctcatcataaga 1263
 Db 1083 CTGACCCCAAGATGTTGGCTTTTGCTTCCAAACTCCCAATCTTATTACCAAGG 1142
 QY 1264 atattcctgaatgggctcaggtgatttcatctgggggtttggatgtctctcaccgtatgt 1323
 Db 1143 AGATTCTTTTACAATCAGTGGACTTTATCTGGCAATGGATGGCCCTCAAGAAAGTTT 1202
 QY 1324 acaatgaatcatccactgtat 1345
 Db 1203 ACAATGACATGTTCTATGAT 1224

RESULT 13

A01157
 LOCUS A01157 1336 bp mRNA PAT 29-JAN-1993
 DEFINITION R. norvegicus mRNA for prelingual lipase protein.
 ACCESSION A01157
 VERSION A01157.1 GI:14748
 KEYWORDS lipase.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1336)
 AUTHORS Williamson, R.
 JOURNAL Patent: WO 8500381-A 3 31-JAN-1985;
 Celltech Ltd
 FEATURES
 Location/Qualifiers
 1..1336
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 7..1194
 /codon_start=1
 /product="prelingual lipase"
 /protein_id="CAA00136.1"
 /db_xref="GI:14749"
 /db_xref="SWISS-PROT:P04634"
 /translation="MWLLITSVISTFGAHLFGKLGPNPEANNISOMITYWYYP
 COEYEVVTEGDIYVYRIPHGKNNSEINIGKRPVYVYLQGLIASATNWIANLPNNLSA
 FMLADAGYDVWLGNSRGNTWKRNVYSPDSVEFWAFSDMAKYVQSLKISFPT
 GQEKIHYHSGOCTTGTAFTSTNPTLAKIKTFALAPVATVYQSPKIKISFPT
 PFKLMFGKKMFLPHTYFDGFLGTGTEVCSREVLDLLCNLTFICGPDKKLNYSREDFV
 VLGNHPCSTVQDFLHWALVRSKGFQAFNWSQSNMLHYNOKTPEVDVSAMTVPV
 AVWNGNDILADPQDVAMLLPKLNLFLHKEILAYNLHDFINWADAPQEVINEMISM
 AED"
 BASE COUNT 409 a 283 c 265 g 379 t
 ORIGIN
 Query Match 22.9%; Score 467.6; DB 9; Length 1336;
 Best Local Similarity 63.5%; Pred. No. 6.8e-107;
 Matches 713; Conservative 0; Mismatches 409; Indels 0; Gaps 0;

QY 244 ctgtgacagcagcattcatgaattagtgaaatcatccaacatcaaggctatccct 303
 Db 80 CTGGAACCTGAAGCAAAATATGAATATTAGTCAGATGAATACTTCTGGGGATATCCAT 139
 QY 304 gtgagaaatgaagtcgcaactgaagatgggtatatacctttctgttaacaggaattcttc 363
 Db 140 GTCAAGAAATGAAGTTGTACTGAAGATGGCTACATTCCTGGGGTCTACAGAATTCCTC 199
 QY 364 gaggcctagtgcaactcaagaagacaggttccaggcgtgtggtgttactcagcagatggcc 423
 Db 200 ATGGGAAGAAATTAATCTGAAATATATAGGCAAGAGACCTGTGTGTATTTCAGCATGGT 259
 QY 424 tagtggaggtgctagcaactgatttccaacctgcccacaataagctggtggttcttc 483
 Db 260 TGAATCCATCAGCACCAAACTGGATTCGAAATCTACCAACACAGCCTGGCCTTTATGC 319
 QY 484 tggcagatgctgttttgacgtgtgagtgagggaacagcaggggaaacgcctgtgtctcgaa 543
 Db 320 TAGCACAGCTGGCTATGATGTGTGGTGGGAACAGTCCAGGAAATATACATGTGCCAGGA 379
 QY 544 aacaaagacactctccatagacacagatgagttctgggtttcagttatgagatgag 603
 Db 380 AAAATGTATACTACTCACCAGACTCAGTTGAATTCCTGGGCTTTTCAGCTTTGATGAAATGG 439
 QY 604 ctgagttgacctctctgagtgataaaactttatttgcagaaaacagcagggaagaaaga 663
 Db 440 CTAATATGACCTTCCCGCCCACTAATACTTCATTGTACAGAAAACCTGGACAGAGAGA 499
 QY 664 tctattatgctgggctattcacaggggaccaccactgggtttattgcattttccaccatgc 723
 Db 500 TACACTATGTTGGTCAATCTCAGGGCACCACTATTGTTTTCATTGCTTCTTACCAATC 559
 QY 724 cagagtggtctcagaaaaatacaaatgtattttcttttagcaccatagccactgttaagc 783
 Db 560 CTACACTGGCCAAAATAATCAAGACGCTTTTATGCTATTAGTCTCCAGTGTACCGTGAAGT 619
 QY 784 atgcaaaagaccccgaggacaaatgtttgtgctgacagatatgatcaagggattgt 843
 Db 620 ATACACAAAGTCCCTTTGAAAAGAAATTCATTTATTCCTACATTCCTTTTCAAGCTTATGT 679

```

Qy 844 ttggcaaaaagaattctctgtatcagaccagagattctcagacaactgtgtattacacctt 903
Dy 680 TTGGCAAGAAATGTTCTCTCCACACACCTACTTTGATGACTTCTTGGTACCGAAGTGT 739
Qy 904 gtggccagggtattctgtatcagattgttagtaataatcatttacttctgtgggtgattca 963
Dy 740 GCTCAGCGGAGGTCTAGATCTCTCTGACGACACATTTATTCATCTCTCTGTGGATTG 799
Qy 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtgcccacactctgtcggaa 1023
Dy 800 ACAAGAAAAAATTAAATGTGAGTCTGTTTGTATGTATCTAGGCGATAATCCAGCAGGA 859
Qy 1024 catctgtcaaaatattctcaactgagccagcagtggaattctgtggaactccgggcat 1083
Dy 860 CATCTGTTCAAGACTTCTCCACTGGGCACAGCTTGTAGATCTGGGAAATTTCAAGCCT 919
Qy 1084 ttgactggggagtgagacaaatactggaataatgcaatcagccaactccctgtaaggt 1143
Dy 920 TCAACTGGGAGGCCATCCAGACATGTTACACTACACACAGAAACGCTCCCTGAAT 979
Qy 1144 acagagtcaagatatgacgtgctccctacagcaatgtggacagaggtcagactggctt 1203
Dy 980 ATGATGTGTGAGCATGACGTGTGCCAGTTGTCAGTGTGGAACGGTGGCAATGACATCCTG 1039
Qy 1204 caaatcagaagaagctgaataatgctgtctgtgaggtgagcaacctcatciaccataaga 1263
Dy 1040 CTGATCCCCAAGATGTGCGCATGCTGCTTCCCAACTCTCCAACTCTCTTCCATTAAG 1099
Qy 1264 atattctgaatgggctcagtgattctcctgggttgggtgctcctcaccogtatgt 1323
Dy 1100 AGATCTTGCTCAACTGACCTGGACTTCATCTGGGCAATGGATGCCCTCAAGAGTTT 1159
Qy 1324 acaatgaataatccatctgtatgagcagcagcagcagcagcagcagcagcagcagc 1365
Dy 1160 ACAATGATGATTTCCATGATGTCAGAGACTAAAAGAATC 1201

RESULT 14
RNLIP 1355 bp mRNA ROD 30-MAR-1995
LOCUS Rat mRNA for lingual lipase.
DEFINITION X02309
ACCESSION X02309
VERSION X02309.1 GI:56595
KEYWORDS glycoprotein; lipase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1355)
Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C.,
Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J.
Molecular cloning and nucleotide sequence of rat lingual lipase
cDNA
Nucleic Acids Res. 13 (6), 1891-1903 (1985)
JOURNAL 8521587
MEDLINE Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.
COMMENT Location/Qualifiers
FEATURES
source
1..1355
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
sig_peptide 49..102
CDS 49..1236
/product="put. signal peptide (aa -18 to -1)"
/feature="lipase precursor"
/codon_start=1
/protein_id="CA26179.1"
/db_xref="GI:56596"
/translaton="MWLLITSVTFGGAGLFGKLPNGNPEANMNISOMITYWGY
COEVVTEGYILGVRIPIPHGKNSENIGKRPVYVLOHGLIASATNWTANLNNSLA
FMLADAGYDWLNSRGNWRSKNVYISPDSEVFASFDEMAKIDLPATINFIQKT
GQEKIHVHVSQGTITGFIATSTNPLAKIKITFYALAPVATVKYITQSPLUKISFPT

```

```

FLFKLMFGKKMELPHYTFDFLGTGTCVSRVLDLQSNLFLIFCGDDKKNLNVSRFDV
YLGHPAGTSVQDFLHMAQLVRSRGFOAFNWSGSPSQMILHYNOKTPEYDVSAMTVPV
AYWNGGNDILADPDQVYAMLLPKLSNLLFKELIAYNHLDFIWAADAPQEVYINEMISM
AED"
mat_peptide 103..1233
/product="lipase (aa 1-377)"
misc_feature 505..507
/note="aa 153 pot. lipid binding site"
misc_feature 1355
/note="polyA site"
BASE COUNT 400 a 291 c 387 t 1 others
ORIGIN
Query Match 22.9%; Score 467.6; DB 95; Length 1355;
Best Local Similarity 63.5%; Pred. NO. 6.8e-107;
Matches 711; Conservative 0; Mismatches 409; Indels 0; Gaps 0;
Qy 244 ctgtgagccagcagcattcatgaattagtgaaatcatccaacatcaaggctatccct 303
Dy 122 CTGCAAAACCTGAAGCAAAATATGAATATTAGTCAGATGATAAATCTACTGGGGATATCCAT 181
Qy 304 gtgggaatatgaagtcgcaactggaagtgggtatatcccttctgttaacagggattctc 363
Dy 182 GTCFAGNATATGAAGTTGTTACTGAAGATGGCTACATCTCGGGGTCTACAGAAATTCCTC 241
Qy 364 gaggcctagtgcacaccttaagaagacaggttccaggcctgtggtgttactcagcagatgcc 423
Dy 242 ATGCGAAGAAATAATCTGAAAAATATAGGCAAGACGCTGTGGTGTATTGTCAGCATGGTT 301
Qy 424 tagtggaggtgctagcaactggtattccaaacctgcccacaaatagctgggtctcattc 483
Dy 302 TGATCGCATCAGCCACAAACTGGATTGCAAAATCTTACCAAAACACAGCGCTGGCCCTTATGC 361
Qy 484 tggcagatgctgttttgacgtgtggatggggaacagcagggggaacagcctgggtctcga 543
Dy 362 TAGCAGACGCTGGCTATGATGTGTGGCTGGGGAACACAGTCAGGAAATACATGGTCCAGGA 421
Qy 544 aacaaagacactctccatagaccagatgagttcttgggtcttcagttatgatgagatgg 603
Dy 422 AAAATGTATACTACTACCAGACTCAGTTGAATTTCTGGCTTTCAGCTTTTATGAAATGG 481
Qy 604 ctagggttgacctctcctgagtgataaacttatttgcagaaaaacggggcaggaagaaaga 663
Dy 482 CTAAATATGACCTTCCCGCCACAAATAAATCTTATTGTACAGAAAACCTGGACAAGAGAAGA 541
Qy 664 tctattatgctggtattctacagggccaccacacacacacacacacacacacacacacac 723
Dy 542 TACACTATGTTGGTCACTTCTCAGGGCACCACTATTGGTTTCATTCGCTTTTCTACCAATC 601
Qy 724 cagagctggctcagaaaaatcaaaatgtatttggcttttagcaccatagccactgttaagc 783
Dy 602 CTACACTGGCCAAAAAATCAAGACGTTTATGATCATAGCTCCAGTTGCTACCGTGAAGT 661
Qy 784 atgcaaaagcccccgggagcaaaatttttggctgacagatatgatgaagggattgt 843
Dy 662 ATACACAAAGTCCCTTGAAAAAGATTTTATTTTCTACATTTCTTTTCAAGCTTATGT 721
Qy 844 ttggcaaaaagaattctctgtatcagcagcagattctcagacaactgtttattacacctt 903
Dy 722 TTGGCAAGAAATATGTTCTGCCCCACACCTACTTGTGATGACTTCTTGGTACCGAAGTGT 781
Qy 904 gtggccagggtattctctgtatcagattgttagtaataatcatttacttctgtgggtgattca 963
Dy 782 GCTCAGCGGAGGTCTAGATCTCTCTGACGACACATTTTATTCATCTCTCTGTGGATTG 841
Qy 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtgcccacactcttgcaggaa 1023
Dy 842 ACAAGAAAAAATTAAATGTGAGTCTGTTTGTATGTGTATCTAGGCGATAATCCAGCAGGA 901
Qy 1024 catctgtcaaaatattctcaactgagcagcagcagcagcagcagcagcagcagcagcagc 1083
Dy 902 CATCTGTTCAAGACTTCTCCACTGGGCACAGCTTGTGTAGATCTGGGAAATTTCAAGCCT 961

```


QY	1142	gtacagagtcagagatagtacggtgccctcacagcaatgtgacaggaaggtcagactggct	1201
Db	993	GTACAACATAAAAAACATCGCTTACCCACATGCTTTGTGGAGCGGGCGGACGGCT	1052
QY	1202	tccaatccagagacggtgaaatgtcctctcgaagtgaccaaacctcatctaccataa	1261
Db	1053	GGCAGACATCATGACATCATCTTTACTGACTCAATCCCCAAGTTGGTGATCACAA	1112
QY	1262	gaattctcctgaatggggtcacgtggatttcattcctgggggttggatgctctctcacggtat	1321
Db	1113	GAACATTCCTGAGTGGGACCATCTGGATTTTATCTGGGGTTGGATGCCCCCTGGAACT	1172
QY	1322	gtacaatgaaatcattccatctcgtatgcagcagaggag	1358
Db	1173	GTATGATGAAATTAATCAGTCTAATGAAGAAATACCAG	1209

Search completed: August 24, 2001, 22:44:11
Job time: 9133 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 20:38:04 ; Search time 97.86 Seconds
(without alignments)
3954.138 Million cell updates/sec

Title: US-09-333-159-45

Perfect score: 2044

Sequence: 1 gtcgaccacgctccggg.....aaaaaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.2	21.7	1137	1 US-08-227-108-2	Sequence 2, Appl
2	444.2	21.7	1137	2 US-09-073-674-2	Sequence 2, Appl
3	444.2	21.7	1140	1 US-08-227-108-4	Sequence 4, Appl
4	444.2	21.7	1140	2 US-09-073-674-4	Sequence 4, Appl
5	444.2	21.7	1146	1 US-08-227-108-6	Sequence 6, Appl
6	444.2	21.7	1146	2 US-09-073-674-6	Sequence 6, Appl
7	444.2	21.7	1528	1 US-08-227-108-1	Sequence 1, Appl
8	444.2	21.7	1528	2 US-09-073-674-1	Sequence 1, Appl
9	48.2	2.4	178	1 US-08-751-782-2	Sequence 2, Appl
10	48.2	2.4	178	2 US-08-925-171-2	Sequence 2, Appl
11	46.4	2.3	43676	3 US-09-356-952-12	Sequence 12, Appl
12	42	2.1	8041	1 US-08-765-081-1	Sequence 1, Appl
13	42	2.1	8041	3 US-09-098-082-1	Sequence 1, Appl
14	42	2.1	8041	5 PCT-US95-06994-1	Sequence 1, Appl
15	40.4	2.0	1140	2 US-08-698-805-5	Sequence 5, Appl
16	38.6	1.9	2815	1 US-08-230-491A-1	Sequence 1, Appl
17	38.6	1.9	2815	2 US-08-619-280A-1	Sequence 1, Appl
18	38.6	1.9	2815	1 US-08-940-391-1	Sequence 1, Appl
19	38.4	1.9	2798	4 US-09-318-448-34	Sequence 34, Appl
20	36.8	1.8	1248	2 US-08-897-340-5	Sequence 5, Appl
21	36.8	1.8	1248	3 US-09-252-329-5	Sequence 5, Appl
22	36.8	1.8	1406	3 US-08-913-842-6	Sequence 6, Appl
23	36.6	1.8	3271	2 US-08-852-806-1	Sequence 1, Appl
24	36.6	1.8	3271	3 US-09-163-669-1	Sequence 1, Appl
25	36.6	1.8	3282	1 US-08-276-852-154	Sequence 154, Appl
26	36.6	1.8	3282	1 US-08-276-852-169	Sequence 169, Appl
27	36.6	1.8	3282	1 US-08-899-575-154	Sequence 154, Appl

C 28	36.6	1.8	3282	1 US-08-899-575-169	Sequence 169, App
29	36.6	1.8	3282	1 US-08-899-575-154	Sequence 154, App
C 30	36.6	1.8	3282	1 US-08-899-575-169	Sequence 169, App
31	36.6	1.8	3282	5 PCT-US95-08743-154	Sequence 154, App
C 32	36.6	1.8	3282	5 PCT-US95-08743-169	Sequence 169, App
33	36.6	1.8	13254	1 US-08-276-852-156	Sequence 156, App
C 34	36.6	1.8	13254	1 US-08-276-852-170	Sequence 170, App
35	36.6	1.8	13254	1 US-08-899-575-156	Sequence 156, App
C 36	36.6	1.8	13254	1 US-08-899-575-170	Sequence 170, App
37	36.6	1.8	13254	1 US-08-899-575-156	Sequence 156, App
C 38	36.6	1.8	13254	1 US-08-899-575-170	Sequence 170, App
39	36.6	1.8	13254	5 PCT-US95-08743-156	Sequence 156, App
C 40	36.6	1.8	13254	5 PCT-US95-08743-170	Sequence 170, App
41	36.4	1.8	234	3 US-08-651-472-89	Sequence 89, Appl
C 42	35.8	1.8	6027	2 US-08-968-542C-1	Sequence 1, Appl
43	35.2	1.7	37895	1 US-08-375-709-1	Sequence 1, Appl
C 44	35.2	1.7	37895	1 US-08-752-929-1	Sequence 1, Appl
45	35.2	1.7	37895	3 US-09-090-793-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-227-108-2
; Sequence 2, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 21.7%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 1.1e-117;
Matches: 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Db 266 AC GCGGGTACGAGCTGTGGTGGGAAACACAGCAGGGGCAACACCTGGCCAGGAGGAATC 325
QY 550 agacactctccatagacacagatgagttctgggttttcagttatgatgagatggctaggt 609
Db 326 TG TACTACTCGCGGAGCTCGCTCGAATCTGGGCTTTTCAGCTTTGACGAGATGCTAAAT 385
QY 610 ttgacctctcagtgatataaacttttttttgagaaaaagggcgagaaaaagatctatt 669
Db 386 ATGACCTTCCCGCCACCATTCAGCTTCATCTTGAAGAAACGGGACAGCAAGCTACACT 445
QY 670 atgctgggtattccagggccaccatgggtttttgcattttccaccatgcccagagc 729
Db 446 AC GTTGGCCATTCCAGGCGACCAACCATTTGTTTCATGCTTTTCACCAATCCCAAGC 505
QY 730 tgggtcagaaaaataatgtattttgcttttagcaccatagccacgtttaaagcatgcaa 789
Db 506 TGGGGAACAGGATCAAAACCTCTCTATGATAGCTCCGTTGCCACCGTGAAGTACACCG 565
QY 790 aaagcccgaggacaaattttttgtgctgcagatatgatgataagggattgtttgca 849
Db 566 AAACCTGTAAACAACTCATGCTCGCTCTCTGTTCTCTTCAAGCTTATATTGGAA 625
QY 850 aaaaagaattctgtatcagaccagatttctcagacacattgtatttaccittttggcc 909
Db 626 ACAAATATTCTACCCACACCACTTCTTGTATCAATTTCTGCCACCGAGTATGCTCCC 685
QY 910 aggtgattctgtatcagattgttagtaataatcatgttactctctgggtggatccaacca 969
Db 686 GCGAGAGGGTGGATCTCTCTGACGAAACGCGCTTTATCATTTGTGGATTGACACTA 745
QY 970 acaatatgaacatgagccgagcaagtgtatatgtctgcccacactctgtggaaacatctg 1029
Db 746 TGAACCTTGACATGAGCTGCTTGGATGTGTATCTGTACATAATTCACGAGGAACATCGG 805
QY 1030 tgcaaaattctctacatgagccagcagtgagtgattctgtgaactccgggcatcttgact 1089
Db 806 TTCAGACGCTGCTCCACTGTGCCAGGCTGTTAAGTCTGGGAGTTCCAGCTTTTGACT 865
QY 1090 gggggagtgagacaaatacttgaaaaatgcaatcgcaactcctgtgaagtgacagag 1149
Db 866 GGGGAAGCCCGAGTTTCAGAAATGATGACATATCATCAGACATGCTCCCTACTACAAC 925
QY 1150 tcagagatgagcgtccctacagcaatgagcagagatgagcagcagcgttttcaaatc 1209
Db 926 TGACAGACATGATGCTCCCAATGCGAGTGGGAACGGTGGCAACGACTTCTGGCCGACC 985
QY 1210 cagaagacgtgaaaatgctctctgagtgagcaccacactctctaccataagaatattc 1269
Db 986 CTCAGATGTTGACCTTTTGTCTTCCAAAGCTCCCAATCTCATTTACCACAGGAATTC 1045
QY 1270 ctgaatgggttcacgtgagtttcatctcctgggttttgagatgctctcaccgtatgacaatg 1329
Db 1046 CTCCTTAAATCACTTTGGACTTATCTGGGCCATGATGATGATGATGATGATGATGATG 1105
QY 1330 aaatcatccatctgatg 1346
Db 1106 AAATGTTTCCATGATG 1122

RESULT 3

US-08-227-108-4
: Sequence 4, Application US/08227108
: Patent No. 5807726
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-Apr-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 21.7%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. 1.1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
QY 250 acccagaagcattcatgaatttagtgaataatcaccacatcccaaggtctatccctgtgag 309
Db 29 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGGATACCCACTGAGG 88
QY 310 aatatgaagtcgcaactgaagatgggtatatctcttctgttaacagagattctcgagcc 369
Db 89 AATATGAAGTTGTGACGGAAGCGGTTATATCTTGGGATCGACAGAAATTCCTTATGGA 148
QY 370 tagtgcacactaagaagacaggttccagcgtgtgtgttactgcacagatggcctagttg 429
Db 149 GGAATAATTCAGAGAAATATAGCGGGAGACCTGTGTGCATTTTTCACACACGGTTTGTCTG 208
QY 430 gaggtgctagcaactggatttcccaacctgcccaactagcctggtggtcttattctggcag 489
Db 209 CATCAGCCCAAACTGGATCTCCAACCTGCCCAACACACGCTTGCCCTTCATCTCGGCCG 268
QY 490 atgctgggttttgacgtgagtgagtgaggggaacagcaggggaaacgctgtctcgaacacaca 549
Db 269 ACCCGGGTACGACGCTGTGGTGGGAACAGCAGGGGCAACACCTGGGCCAGGAGGAATC 328
QY 550 agacactcccatagacacaaatgagttctgggttttcagtttatgatgagatggctaggt 609
Db 329 TGTACTACTCGCCCGGACTCGTCTGAATTTCTGGGCTTTTCAGCTTTGACGAGATGCTAAAT 388
QY 610 ttgaccttctcgtcagtgataaacttttttgcagaaaaacggcgagaaaaagatctatt 669
Db 389 ATGACCTTCCCGCCACCATTCAGCTTCATCTTGAAGAAACGGGACAGCAAGCTACACT 448
QY 670 atgtcggctattccaggggcaccacatgggttttatgtcattttccaccatgcccagagc 729
Db 449 AC GTTGGCCATTCCAGGCGACCAACCATTTGTTTCATGCTCTTTTCCACCAATCCCAAGC 508
QY 730 tggctcagaaaaataaataatgtattttgtcttagcaccatagccactgttaagcatgcaa 789
Db 509 TGGCGAAGCGGATCAAAACCTTCTATGATCTAGCTCCGTTGCCACCGTGAAGTACACCG 568

QY 790 aagcccccggagcaaaatttttggctgcccagatgatgatcaaggaggtgtttggca 849
Db 569 AAACCCCTGTTAAACAAATCATCTGCTGCTCCCTGCTTCAAGCTTATATTTGGA 628
QY 850 aaaaagaattctgtatcagaccagattctcagacaactgtttattacatttggccc 909
Db 629 ACAAAATATTCTACCAACACATCTTTGTATCAATTTCTGCGCACCGAGGTATGCTCCC 688
QY 910 aggtgattctgtacagattgttagtaataatctacttctggtggattcaacacca 969
Db 689 GCGAGACGGTGAATCTCTCTGACAGACGCTGTTATCATCTTTGGATTGACACTA 748
QY 970 acaatatgaacatgagccgagcaagtgtatatgtgcccacactcttgcgtggaacattcg. 1029
Db 749 TCAACTTGAACATGAGTGGCTGGATGTATCTGTACATAATCCAGCAAGCAATCGG 808
QY 1030 tgcataattctacactgagccagccagcagtggaattctgtggaactccgggcatttgact 1089
Db 809 TTCAGACGCTCTCCACTGGTCCGAGGCTGTTAAGTCTGGGAAGTTCAGGCTTTGACT 868
QY 1090 gggggagtgagacaaatctggaataatgcaatcagcccaactcctgtaagggtacagag 1149
Db 869 GGGGAAGCCAGTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAGC 928
QY 1150 taagatatgacgctccctacagaaatgagcagaggtgagagctggtttcaaatc 1209
Db 929 TCACAGACATCATGTGCAATCGCAGTGTGGAACGGTGGCAAGCACTTGTGCGCCGACC 988
QY 1210 cagaagcgtgaaatcgtctctgagtgagtgacaaacctcatcaccataagaatttc 1269
Db 989 CTCAGAGTGTGACCTTTTGTCTTCCAGAGTCCCAATCTCATTTACACAGGAAGATTC 1048
QY 1270 ctgaatgggctcagctgattctcctgggtttggatgctcctcaccgtatgacaatg 1329
Db 1049 CTCCTTACATCACTTGGACTTTATCTGGCCATGGATGCCCTCAAGCGTTTACAAATG 1108
QY 1330 aaateatccatctgatg 1346
Db 1109 AAATGTTTCCATGATG 1125

RESULT 4

US-09-073-674-4
; Sequence 4, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073.674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-09-073-674-4

Query Match 21.7%; Score 444.2; DB 2; Length 1140;
Best Local Similarity 62.8%; Pred. No. 1..1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 acccagaagcattcatgaattattagtgaaatcatccaacatcaaggctcatccctgtgagg 309
Db 29 ACCGTGAAGTCACCATGAATATAGTCAGATGATCACTACTGGGATACCCAGCTGAGG 88
QY 310 aatatgaagtcgonaactgaagatgggtatatcccttctgttaacaggagattcctcgagcc 369
Db 89 AATATGAAGTTGTGACCGAAGACGTTATATCTTGGGATCGACAGAAATCTCTTATGGGA 148
QY 370 tagtgcaacctgaagaacagaggttccagcctgtgggttactcagcagatgacctagt 429
Db 149 GGAAATAATCAGAGATATAGGCGGAGACCTGTGCAATTTTGCACACAGGTTGCTCG 208
QY 430 gagtgcctagcaactgattcccaacctgcccacaatagcctgggcttctcattctggcag 489
Db 209 CATCAGGCCAANAATGATCTCCAACTGCGCCCAACACAGCGCTGGCCTTCATCTCTGGCCG 268
QY 490 atgctggttttgacgtgtggtgggaacagcaggggaaacagcctgggtctcgaaaaacaca 549
Db 269 ACGCGGGTACGACGCTGTGGCTGGGAACAGCAGGGGCAACACCTGGGCCAGGAGGAATC 328
QY 550 agacactctccatagacacaagatgagttctgggttttaagtattatgagatgagctaggt 609
Db 329 TGTACTACTCGCCGACTCGCTCGAATTTCTGGGCTTTTTCAGCTTTTACAGAGATGGCTAAT 388
QY 610 ttgacctctcgtcagtgataaactttatttgcagaaaaacggccagagaaagatcattt 669
Db 389 ATGACCTTCCCGCCACCATTCATCTTTGAAGAAAAACGGGACAGACAGCTACTACT 448
QY 670 atgtcggctattcagagggccaccaccatgggctttattgcattttccaccatgcccagagc 729
Db 449 ACGTGGCCATTCCAGGGCACCACTTGGTTTTCATCGCTTTTCCACCAATCCCAAGC 508
QY 730 tggctcagaaaaatcaaaatgatttctttagcaccatagccactgttaagatgcaa 789
Db 509 TGGGAAACGGGATCAAAACCTTCTATGCAATTTAGTCTCGGTTGCGCACCGTACACCG 568
QY 790 aaacccccgggaccaaattttgtgctgcccagatgatgatcaaggagattgtttggca 849
Db 569 AAACCTGTTAAACAAATCATGCTGCTGCTCTTCTCTTCAAGCTTATATTGGA 628
QY 850 aaaaagaattctgtatcagaccagattctcagacaactgtttattacatttggccc 909
Db 629 ACAAAATATTCTACCAACACATCTTCTTGTATCAATTTCTGCGCACCGAGGTATGCTCCC 688
QY 910 aggtgattctgtacagattgttagtaataatcatttacttctggtggattcaacacca 969
Db 689 GCGAGCGGTGATCTCTCTGACAGCAAGCCCTGTTTATCATTTGTGGATTGTGACACTA 748
QY 970 acaatatgaacatgagccgagcaagtgtatatgtgcccacactcttctgtggaacattcg 1029
Db 749 TGAACCTTGAACATGAGTGGCTTGGATGTATCTGTACATAATCCAGCAAGCAATCGG 808
QY 1030 tgcataattctacactgagccagccagcagtggaattctgtggaactccgggcatttgact 1089

Db 809 TTCAGACGCTGCTCCACTGGTCCAGGCTGTAACTCTGGGAGTTCCAGCTTTTGACT 868
Qy 1090 gggggagtgagaccacaaatctggaaaaatgcaatcagccaaactcctgtgaaggtagag 1149
Db 869 GGGGAAGCCCACTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAACC 928
Qy 1150 tcagagatatgacgtccctcagcaaatgtggacagagggtcaggactggctttcaaatc 1209
Db 929 TGACAGACATGATGTCATCGCAATGCGAGTGTGGAACGGTGGCAACACATGCTGGCCGACC 988
Qy 1210 cagaagacgtgaaaatgctctctgctgagtgagtgagcaaacctcatctaccataagaatattc 1269
Db 989 CTCAGATGTTGACCTTTTGCTTTCCAGCTCCCAATCTCATTTACACAGGAGATTC 1048
Qy 1270 ctgaatgggctcactgattcatctctgggtttggatgctcctcactgataatg 1329
Db 1049 CTCCTTACAATCATTGGACTTTATCTGGGCATGGATGCCCTCAAGCGGTTTACAATG 1108
Qy 1330 aaatcatccatctgatg 1346
Db 1109 AAATGTTTCCATGATG 1125

RESULT 5

US-08-227-108-6

; Sequence 6, Application US/08227108

; Patent No. 5807726

; GENERAL INFORMATION:

; APPLICANT: Blanchard, Claire

; APPLICANT: Benicourt, Claude

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,108

; FILING DATE: 03-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fanucci, Allan A.

; REGISTRATION NUMBER: 30,256

; REFERENCE/DOCKET NUMBER: 7620-033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-5090

; TELEFAX: 212 869-8864/9741

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1146 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; US-08-227-108-6

Query Match

Best Local Similarity 21.7%; Score 444.2; DB 1; Length 1146;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 250 acccagaagcattcatgaattagtgaaatcaccacatcaaggctatccctgtgagg 309

Db 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGGATACCCAGGTGAGG 85
Qy 310 aatatgaagtgcgaactgaagatgggtatatactcttctgttaacaggattctctcagggcc 369
Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTTTGGGATCCAGAGAAATTCCTTTATGGA 145
Qy 370 tagtgaaccttaagaagacaggttccagggcctgtgtgttactcagcatggccttagttg 429
Db 146 GGAATAATTCAGAGAAATATAGCCCGGAGACCTGTTCATTTTGGCAACACGGTTTGCTCG 205
Qy 430 gagggtcagcaactggatttccaaactgcacaaatagcctgggttccattctctgagag 489
Db 206 CATCAGCCACAACTGGGATCTCCAACCTGCCCAACAGAGCTGGCTTCATCTCTGGCCG 265
Qy 490 atgctgggtttgacgtgtggatggggagacagcaggggaaacacccctggctcgaacacaca 549
Db 266 ACGCCGGGTACGACGTGTGGCTGGGGAACAGCAGGGGCAACACCTGGGCCAGGAGATC 325
Qy 550 agaacctctccatagaccaagatgagttctctgggttccagtttatgatgagatggctaggt 609
Db 326 TGTACTACTCGCCCGACTCCCGTCGAATTTCTGGCTTTTTCAGCTTTGACGAGATGGCTAAAT 385
Qy 610 ttgaccttctcagtgatgataaactttatttgacagaaacagggccaggaagaaagatctatt 569
Db 386 ATGACCTTCCCGCCACCATTTGACTTCATCTTTGAAGAAAACGGGACAGGACTACACT 445
Qy 670 atgctgggtattcacagggccaccacacacacacacacacacacacacacacacacacac 729
Db 446 ACGTTGGCCATTCCAGGGGACACCATTTGTTTCATCGCTTTTCCACCAATCCCAACG 505
Qy 730 tgggtcagaaaaatcaaatgtattttgttttagcaccacacacacacacacacacacacac 789
Db 506 TGGGAAACGGATCAAAACCTTCTATGATTTAGCTCCGTTGCCACCGTGAAGTACACCG 565
Qy 750 aagaccccgggacacaaattttgtgtgcagacacacacacacacacacacacacacacac 849
Db 566 AAACCTCTTAAACAACTCATGTGCTGCTCCCTTCTCTCAAGCTTATATTGGAA 625
Qy 850 aaaaagaatttctgtatcagaccagatttctcagacacacacacacacacacacacacac 909
Db 626 ACAATAATTCTACCCACACCATTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCC 585
Qy 910 agg-gattctctgacagattgttagtaataatcttactctctgggtggatcaacacca 969
Db 686 GCGAGACGTTGGATCTCTCTGACGACGCGCTTTTATCATTTTGTGGATTGACACTA 745
Qy 970 acatataaacatgagccgagcaagtgtatatgtgcacacacacacacacacacacacacac 1029
Db 746 TGAACATGATGATGCTGGATGTGTATCTCTCATATATCCAGCAGGAACATCGG 805
Qy 1030 tgcnaaatattctacactggagccagggcagtgaaatctcgtgaaactccgggctatttgact 1089
Db 806 TTCAGAACGCTGCTCCACCTGGTCCAGGCTGTTAACTCTGGGAAGTTCCAACTTTTGACT 865
Qy 1090 gggggagtgagac 1149
Db 866 GGGGAAGCCCACTTCAGAACATGATGCACTATCATCAGAGCATGCTCTCTACTACAAC 925
Qy 1150 tcagagatatgacggctccctcagcaaatgtggacaggggtcaggactggcttcaaatc 1209
Db 926 TGACAGACATGATGTGCAATTCGCAATGCGAGTGTGGAACGGTGGCAACGACTTGTCTGGCCGACC 985
Qy 1210 cagaagacgtgaaaatgctctctcagggagacacacacacacacacacacacacacacac 1269
Db 986 CTCAGATGTTGACCTTTTGCTTTTCCAAAGTCCCAATCTCATTACCACAGGAAGATTC 1045
Qy 1270 ctgaatgggtcagtgatttcatctctgggtttggatgctcctcactcagcgtatgtcaatg 1329
Db 1046 CTCCTTACAATCATTGGACTTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATG 1105
Qy 1330 aaatcatccatctgatg 1346

Db 1106 AAATGTTTCATGATG 1122

RESULT 6
US-09-073-674-6
; Sequence 6, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-6

Query Match 21.7%; Score 444.2; DB 2; Length 1146;
Best Local Similarity 52.8%; Pred. No. 1.1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 accagaagcattcgaatatttagtgaatcaccacatcccaagcgtatccctgtgag 309
|||||
Db 26 ACCCTGAAGTGACCATGAATATAGTCAGATGATCACCCTACTGGGGATACCCAGCTGAGG 85
|||||
QY 310 aatatgaagtcgaactgaagtgaggtatattcttctgttaacaggattctcagagcc 369
|||||
Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTTGGGATCGACAGAATTCCTTATGGGA 145
|||||
QY 370 tagtgcacctaagaagcaggttccagcgtgtgttactcagcatggcctagtgtg 429
|||||
Db 146 GGAATAATTCAGAGATATAGGCGGAGACCTGTGTGATTTTTCACACGCGTTTGTCTCG 205
|||||
QY 430 gaggtgtagcaactggtattccaaacctgcccacaaatagcttggttcttcttcagcag 489
|||||
Db 206 CATCAGCCACAACTGGATCTCCAACTGCCCACACACAGCTGCGCTTCATCTGCGCG 265
|||||
QY 490 atgctggttttgcagtggtgagcagcaggggaaacgcctgtctcgaacacaca 549
|||||
Db 266 ACGCCGGGTACGACGCTGTGCTGGGAAACAGCAGGGGCAACACCTGGCGCAGGAGATC 325
|||||
QY 550 agacactctccatagaccagaatgagttcttggtcttcttcagttatgatcagatggctagg 609
|||||

Db 326 TGTACTACTCGCCGACCTCCGTCGAATTCCTGGGCTTTTACGCTTTTGACGAGATGGCTAAAT 385

QY 610 ttgaccttctcgtcagtgatataaaactttattttgcagaaaaacggccagggaaaaaatctctatt 669
|||||
Db 386 ATGACCTTTCCCGCCACCATTTGACTTTCATCTTTGAAGAAAACGGGACAGGACAGCTACACT 445
|||||
QY 670 atgtcggctattcacagggccacaccattggctttattgtattttccattttccaccatccagagc 729
|||||
Db 446 AGCTTGGCCATTCCAGGGCACCACCATTTGGTTTCATCGGCTTTTCACCAATATCCCAAGC 505
|||||
QY 730 tggctcagaaaaatcaaaatgtattttgttttagcaccatagccactgtttaagcatgcaa 789
|||||
Db 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGCTCCCGTTCACCGTGAAGTACACCG 565
|||||
QY 790 aaagcccggggacaaaattttgtgtcctccagatgatgatgaaggggatttggttggca 849
|||||
Db 566 AAACCTTGTAAACAAACATCATGCTGCTGCTCTCTTCAAGCTTATATTGGA 625
|||||
QY 850 aaaaagaatttctgtatcagaccagatttctcagacaacttcttatttcttctgtgccc 905
|||||
Db 626 ACAAAATATTTACCCACACACACTTCTTTGATCAATTTCTGGCCACGGAGTATGCTCCC 685
|||||
QY 910 aggtgattcttgatcagattgttagtaataatcatgttactcttgggtgggattcaacacca 969
|||||
Db 686 GCGAGACGGTGGATCTCTCTCGAGCAACGCCCTGTTTATCATTTGTGGATTTCACACTA 745
|||||
QY 970 acaatatgaacatgagccgagcaagtgtatgtctgctccacacactcttctgtggaaactctg 1029
|||||
Db 746 TGAACCTTGAACATGAGTCGCTTGGATGTGTATCTGTACATAATCCAGCAGGAACATCGG 805
|||||
QY 1030 tgcataatattctacactggagccagggcagtgatgaattctgtgactccgggcttctgact 1089
|||||
Db 806 TTCACAAAGCTGCTCCACTGGTCCCAGGCTGTTAAAGTCTGGGAAGTTCCAAAGCTTTTGACT 865
|||||
QY 1090 gggggagtgagaccacaaaatctggaaaaatgcaatcagccaactcctgttaagttacagag 1149
|||||
Db 866 GGGGAAGCCCGAGTTTCAGAAACATGATGCACATATCATCAGAGCATGCTCCTACTACAAC 925
|||||
QY 1150 tcagagatgatcaggttccctcagcaaatgtgacagagggtcagggactggtcttcaaatc 1209
|||||
Db 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTCTCTGGCGGAC 985
|||||
QY 1210 cagangacgtgaaaaatgctgctctctgagggtgacaaacctcatctaccataagaatttc 1289
|||||
Db 986 CTCAGATGTTGACCTTTTGCTTTCCAAAGCTCCCAAGCTCCCAATCTCATTTACCAGGAAGATC 1045
|||||
QY 1270 ctgaatgggtcagcgtgatttccatctctgggtttggtgactctcctaccgtatgacaatg 1329
|||||
Db 1046 CTCCCTACATACATTTGGACTTTATCTGGCCATGATGCTCCCTCAAGCGGTTTACAATG 1105
|||||
QY 1330 aaatcatcctctgatg 1346
|||||
Db 1106 AAATGTTTCATGATG 1122

RESULT 7
US-08-227-108-1
; Sequence 1, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:

Db 1106 AAATGTTTCATGATG 1122

RESULT 6
US-09-073-674-6
; Sequence 6, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-09-073-674-6

Query Match 21.7%; Score 444.2; DB 2; Length 1146;
Best Local Similarity 52.8%; Pred. No. 1.1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 accagaagcattcgaatatttagtgaatcaccacatcccaagcgtatccctgtgag 309
|||||
Db 26 ACCCTGAAGTGACCATGAATATAGTCAGATGATCACCCTACTGGGGATACCCAGCTGAG 85
|||||
QY 310 aatatgaagtcgaactgaagtgaggtatattcttctgttaacaggattctcagagcc 369
|||||
Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTTGGGATCGACAGAATTCCTTATGGGA 145
|||||
QY 370 tagtgcacctaagaagcaggttccagcgtgtgttactcagcagcagcaggtttg 429
|||||
Db 146 GGAATAATTCAGAGATATAGGCGGAGACCTGTGTGATTTTTCACACGCGTTTGTCTG 205
|||||
QY 430 gaggtgtcgaactggtattccaaacctgcccacaaatagcttgccttcttctcagcag 489
|||||
Db 206 CATCAGCCACAACTGGATCTCCAACTGCCCACACACAGCTGCGCTTCATCTGCGCG 265
|||||
QY 490 atgctggttttgcagtcgtgtgagtgagcagcagcagcagcagcagcagcagcagc 549
|||||
Db 266 ACGCCGGGTACGACGCTGTGCTGGGAAACAGCAGGGGCAACACCTGGCGCAGGAGATC 325
|||||
QY 550 agacactctccatagaccagaatgagttcttggtcttcttcttcttcttcttcttctt 609
|||||

Db 326 TGTACTACTCGCCGACCTCCGTCGAATTCCTGGGCTTTTACGCTTTTGACGAGATGCCTAAAT 385

QY 610 ttgaccttctcgtcagtgatataaactttattttgcagaaaaacggccagggcagaaatctctatt 669
|||||
Db 386 ATGACCTTTCCCGCCACCATTTGACTTTCATCTTTGAAGAAAACGGGACAGGATACACT 445
|||||
QY 670 atgtcggctattcacagggccacaccattggttttatttgcattttccaccattccagagc 729
|||||
Db 446 AGCTTGGCCATTCCAGGGCACCACCATTTGGTTTCATCGGCTTTTCACCAATCCCAAGC 505
|||||
QY 730 tggctcagaaaaatcaaaatgtattttgtttagcaccatagcactgtttaagcatgcaa 789
|||||
Db 506 TGGCGAAAACGGATCAAAACCTTCTATGCAATAGCTCCCGTTCACCGTGAAGTACACCG 565
|||||
QY 790 aaagcccgaggacaaaattttgtgtcgtccagatgatgatgaaggggatttggttgcca 849
|||||
Db 566 AAACCTTGTAAACAAACATCATGCTGCTGCTCTCTTCAAGCTTATATTGGA 625
|||||
QY 850 aaaaagaatttctgtatcagaccagatttctcagacaacttcttatttcttcttcttcttctt 905
|||||
Db 626 ACAAAATATTTACCCACACACACTTCTTTGATCAATTTCTGGCCACGGAGTATGCTCCC 685
|||||
QY 910 aggtgattctgtatcagattgttagtaataatcatgttactcttcttcttcttcttcttctt 969
|||||
Db 686 GCGAGACGGTGGATCTCTCTCGAGCAACGCCCTGTTTATCATTTGTGGATTTCACACTA 745
|||||
QY 970 acaatatgaacatgagccgagcaagtgtatgtctgtcgtccacacactcttcttcttcttctt 1029
|||||
Db 746 TGAACCTTGAACATGAGTCGCTTGGATGTGTATCTGTACATATATCCAGCAGGAACATCGG 805
|||||
QY 1030 tgcataatattctacactggagccagggcagtgatgaattctgtgactccggcgttcttctt 1089
|||||
Db 806 TTCAGAACGCTGCTCCACTGGTCCCAGGCTGTTAACTCTGGGAAGTTCCAAAGCTTTTGA 865
|||||
QY 1090 gggggagtgagacacaaaatctggaaaaatgcaatcagcacaactcttcttcttcttcttct 1149
|||||
Db 866 GGGGAAGCCAGTTTCAGAACATGATGCACATATCATCAGAGCATGCCTCTACTACAAC 925
|||||
QY 1150 tcagagatgatcaggttccctcagcaaatgtgacagagggtcagggcttcttcttcttctt 1209
|||||
Db 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTCTCTGGCGGAC 985
|||||
QY 1210 cagangacgtgaaaaatgctctctctcgtgaggtgacaaacctcatctaccataagaatttc 1289
|||||
Db 986 CTCAGATGTTGACCTTTTGCTTTCCAACTCCCAACTCTCATTTACCAGAGAGATTC 1045
|||||
QY 1270 ctgaatgggtcagcgtgatttctctctctctctctctctctctctctctctctctctct 1329
|||||
Db 1046 CTCCCTACATACATTTGGACTTTATCTGGCCATGATGCTCCCTCAAGCGGTTTACAATG 1105
|||||
QY 1330 aaatcatcctctgag 1346
|||||
Db 1106 AAATGTTTCATGATG 1122

RESULT 7
US-08-227-108-1
; Sequence 1, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:

MEDIUM TYPE:	Floppy disk
COMPUTER:	IBM PC compatible
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/08/227,108
FILING DATE:	03-APR-1994
CLASSIFICATION:	435
ATTORNEY/AGENT INFORMATION:	
NAME:	Fanucci, Allan A.
REGISTRATION NUMBER:	30,256
REFERENCE/DOCKET NUMBER:	7620-033
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	212 790-9090
TELEFAX:	212 869-8864/9741
TELEX:	66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:	
SEQUENCE CHARACTERISTICS:	
LENGTH:	1528 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	double
TOPOLOGY:	linear
MOLECULE TYPE:	CDNA TO mRNA
US-08-227-108-1	
Query Match	21.7%
Best Local Similarity	62.8%
Matches	689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
Query	250 acccagaagcattcaatgtaattagtgaatacatcccaacatcaaggcttcgtggag 309

RESULT: 8
US-09-073-674-1
; Sequence 1, Application US/09073674
; Patent No. 598189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA to mRNA
 US-09-073-674-1

Query Match	21.7%	Score 444.2	DB 2	Length 1528
Best Local Similarity	62.8%	Pred. No. 1.3e-117		
Matches 689	Conservative 0	Mismatches 408	Indels 0	Gaps 0
Qy 250	accccaagaatcctgaatattagtagaatacctcccaacatcaagctatccctctgagg	309		
Db				
26	ACCTGAAGTGACCATGAATATAAGTCAGATGATCACCTACTGGGATACCCAGCTGAGG	85		
Qy 310	aatagaagtcgaactgaagatgggtatataccttctctgttaacaggaattcctcaggcc	369		
Db				
86	AATATGAAGTTGTGACCGAAGACGGTTATATCTTGGGATCGACAGAAATCTCTTATGGGA	145		
Qy 370	tagtgcacctaaagaacagaggttccaggcctgtgttgtaactgcagcatggcctagtgtg	429		
Db				
146	GGAAAAATTCAGAGATATAGCCCGAGACCTGTGTGATTTTGCACACGGTTTCTCTCG	205		
Qy 430	gggtgtcagcaactggatttccaaactgcgaacaaatagcctgggcttcatctctgcag	489		
Db				
206	CATCAGCCACAACATGGATCTCCAACTGCCCAACAACAGCCTGGCTTCATCTTGGCCG	265		
Qy 490	atgctggttttgacgtgtggatggggaacagcagggggaacgcctggtctcgaaaaacaa	549		
Db				
266	ACGCCGGGTACAGCCTGTGGCTGGGGAACACAGAGGGCAACACCTGGCCGAGGGAATC	325		
Qy 550	agacactctccatagacaaatagttctggcttccagtttccagttatgatagatgctagct	609		
Db				
326	TGTACTACTTCGCCGACCTCCGTCGTAATCTCTGGGCTTTACGTTTGACGATGGCTTAAT	385		
Qy 610	ttgacctctctgcagtataaactttatttgcgaaaaacggcggaagaaagatctatt	669		
Db				
386	ATCACCTTCCGCCACCATTTGACTTCATCTTGAAGAAACGGGACAGGACAAGCTACACT	445		
Qy 670	atgtcggtattccagggcaccacattggctttattgatttccacatgccagagc	729		
Db				
446	ACGTTGGCCATTTCCAGGGCACACCATTTGTTTCATCGCTTTTCACCAGTCCCAAGC	505		
Qy 730	tggctcagaaaaatcaaaatgtatttggctttagcaccatagccactgttaagcatgcaa	789		
Db				
506	TGCGAAACGGATCAAAACCTTCTATGATAGTCCCGTTGCCACCGTGAAGTACACCG	565		
Qy 790	aaagcccccgggaccaaaatttggctgctcagatatgatatgaagggattgtttggca	849		
Db				
566	AAACCTGTGTTAAACAACATCATGCTGTCGCCCTTGCTTCTCTTCAAGCTTATATTGGAA	625		
Qy 850	aaaaagaattctgtatcagaccagattctcagacaaactgttatttacctctgtggcc	909		
Db				
626	ACAAATATTTACCCACACACACTTCTTTGATCAATTTCTCGCCACCGAGGTATGCTCC	685		
Qy 910	aggtgattcttgatcagattttagtaatacatgattacttctgggtggattccaaccca	969		
Db				
686	CGGAGACGGTGGATCTCTCTCGACCAACGCCCTGTTTATCATTTGTGGATTGACACTA	745		
Qy 970	acaatatgaacatgagccagcaagtgatatgtgccccacactctctgtcgggaacatctg	1029		
Db				
746	TGAACCTTGACATGAGTCTGCTTGGATGTGTATCTGTACATTAATCCACAGGAAACATCGG	805		
Qy 1030	tgcaaaatattctacactgagccaggcagtgtaattctgtgtaactccgggcatcttgact	1089		
Db				
806	TTCAGAACGTGCTTCCACTGGTCCCAGGCTGTAAAGTCTGGGAAGTTTCCAAGCTTTTGACT	865		
Qy 1090	ggggggagtgagacaaaaatctggaaaaatcgaaatcgaccacactcctctgtaaggtacag	1149		
Db				
866	GGGGAAGCCCGATTACAGACATGATGCACTATATCATCAGAGCATGCCCTCCCTACTACAAC	925		
Qy 1150	tcagagatatgacggttccctacagcaatgtggacagagggtcaggactggtcttcaaatc	1209		
Db				
926	TGACAGACATGATGTGCCAAATCGCAGTGTGGAAACGGTGGCAACGACTTCTGCTGGCGAAC	985		

Qy	1210	cagaagacgtgaaatactgctctctctgagtgagacacacccatctacataagaatattc	1269
Db	986	CTCAGATGTTGACCTTTTGGCTTTCCAAAGCTCCCCAACTCTATTACCACAGGAAGATTC	1045
Qy	1270	ctgaatgggctcacgtgattctctgggtttggatgctcctcacccgtatgtacaaatg	1329
Db	1046	CTCCTTACAACTCTGGACCTTTATCTGGGCCATGGATGCCCTCAACGGCGTTTACAATG	1105
Qy	1330	aaatcattccatctgtatg	1346
Db	1106	AAATGTGTTCCATGATG	1122
RESULT 9			
US-08-751-782-2			
; Sequence 2, Application US/08751782			
; Patent No. 5921352			
; GENERAL INFORMATION:			
; APPLICANT: Helntz, Nathaniel			
; APPLICANT: Gubbay, Johnathan			
; APPLICANT: Skinner, Michael			
; TITLE OF INVENTION: A CDNA Library Prepared during			
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof			
; NUMBER OF SEQUENCES: 15			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: David A. Jackson, Esq.			
; STREET: 411 Hackensack Ave, Continental Plaza, 4th			
; STREET: Floor			
; CITY: Hackensack			
; STATE: New Jersey			
; COUNTRY: USA			
; ZIP: 07601			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC Compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/751,782			
; FILING DATE: 18-NOV-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Jackson Esq., David A.			
; REGISTRATION NUMBER: 26,742			
; REFERENCE/DOCKET NUMBER: 600-1-190			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 201-487-5800			
; TELEFAX: 201-343-1684			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 178 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: cDNA			
; HYPOTHETICAL: NO			
; IMMEDIATE SOURCE:			
; CLONE: 10.2			
; US-08-751-782-2			
Query Match 2.4%; Score 48.2; DB 1; Length 178;			
Best Local Similarity 55.8%; Pred. No. 0.00013;			
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0			
Qy	541	gaaacacaaagacactctccatagaccaaagatgagttctgggtcttcagttatgatgaga	600
Db	8	GGAACATGTCGGCCCTAGACCAGATCTTAAGAATTTGGGATTTTACTTTTAAATGAC	67
Qy	601	tggtcagggttgacctctcctgcagtgataaaactttatttgcagaaaaacggccagggaaa	660
Db	68	AAATAGAATACACGCTCCGACCATCTATTTCCTGAATGAACAAGAACACAC	127

ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44Mb storage
COMPUTER: IBM PC compatible/Pentium
OPERATING SYSTEM: MS-Windows 3.1
SOFTWARE: Word for Windows-6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,081
FILING DATE: March 26, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06994
FILING DATE: June 7, 1995
APPLICATION NUMBER: US 08/265,714
FILING DATE: June 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: CHOR-1-10286
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-682-8100; 1-206-224-0718 (direct)
TELEFAX: 1-206-224-0779
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli O157:H7
STRAIN: 86-24 NALR
IMMEDIATE SOURCE:
CLONE: PEAR
US-08-765-081-1

Query Match 2.1%; Score 42; DB 1; Length 8041;
Best Local Similarity 53.4%; Pred. No. 0.058;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 33 aaaaatgtgaagagtttttaaaacccacacaaattcttcttacttttagaattagttgttaca 92
Db 1541 AAAGTATGTGAAGCTTTATTTTAAACACACAGATAATTGATTTTAAAAATAATGTTAGA 1482
QY 93 ttggcagggaataaaatgcagatgttgaccatgttgaaacctgtgcaagacagtg 152
Db 1481 TTCCAGATAATTTATAAGAAATACCGACATTAACAGGTAAGTCGTATGGCTATGG 1422
QY 153 atgtctcacagaatggaatgtggtctctgattctgtg 195
Db 1421 ATATGNCATTTGTTTAAATGTTAAATGTTGTTGTTGG 1379

RESULT 13
US-09-098-082-1/c
Sequence 1, Application US/09098082
Patent No. 6040421
GENERAL INFORMATION:
APPLICANT: Tarr, P.I., Bilge, S.S., Besser, T.E., Vary Jr., J.C.
TITLE OF INVENTION: Escherichia Coli O157:H7 Epithelial Adhesin
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44Mb storage
COMPUTER: IBM PC compatible/Pentium II
OPERATING SYSTEM: MS-Windows 95
SOFTWARE: Word for Windows-6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,082
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/765,081
FILING DATE: March 26, 1997
APPLICATION NUMBER: PCT/US95/06994
FILING DATE: June 7, 1995
APPLICATION NUMBER: US 08/265,714
FILING DATE: June 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Shelness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: CHOR-1-12402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-682-8100; 1-206-224-0735 (direct)
TELEFAX: 1-206-224-0779
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli O157:H7
STRAIN: 86-24 NALR
IMMEDIATE SOURCE:
CLONE: PEAR
US-09-098-082-1

Query Match 2.1%; Score 42; DB 3; Length 8041;
Best Local Similarity 53.4%; Pred. No. 0.058;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 33 aaaaatgtgaagagtttttaaaacccacacaaattcttcttacttttagaattagttgttaca 92
Db 1541 AAAGTATGTGAAGCTTTATTTTAAACACACAGATAATTGATTTTAAAAATAATGTTAGA 1482
QY 93 ttggcagggaataaaatgcagatgttgaccatgttgaaacctgtgcaagacagtg 152
Db 1481 TTCCAGATAATTTATAAGAAATACCGACATTAACAGGTAAGTCGTATGGCTATGG 1422
QY 153 atgtctcacagaatggaatgtggtctctgattctgtg 195
Db 1421 ATATGNCATTTGTTTAAATGTTAAATGTTGTTGTTGG 1379

RESULT 14
PCT-US95-06994-1/c
Sequence 1, Application PC/TUS9506994
GENERAL INFORMATION:
APPLICANT: Children's Hospital & Medical Center
APPLICANT: University of Washington
APPLICANT: Washington State University Research Foundation
APPLICANT: TARR, PHILLIP I
APPLICANT: BILGE, SIMA S
APPLICANT: BESSER, THOMAS E
APPLICANT: VARY JR, JAMES C
TITLE OF INVENTION: ESCHERICHIA COLI O157:H7 EPITHELIAL ADHESIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON AND KINDNESSPLLC

Query Match 2.0%; Score 40.4; DB.2; Length 1140;
Best Local Similarity 45.5%; Pred. No. 0.059;
Matches 143: Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Search completed: August 24, 2001, 22:38:56
Job time: 7252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 19:50:13 ; Search time 2518.61 Seconds
(without alignments)

Title: US-09-333-159-45
Perfect score: 2044
Sequence: 1 gtcagccacgcgtccggg.....aaaaaaaaaagggccgccc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```



```
/dev_stage=0 day neonate"
153.662
/notes="putative"
/codon_start=1
/protein_id="BAB31766.1"
/db_xref="GI:12859755"
/translacion="MSEILSRVMTVSHRVEIWLILVAYLLQNVNSGHLPTKAAADPE
AFMNVSEIHKRGYFSEYEVATEDGYILSVNRIPRGOTRLKESRPPVLLQHLGLG
DASMTISLNPNSLGFILADAGFDVMMGNSRGNTWSRKHKTLISIDQDFWAFRYKNTL
RVEVDTPMP"
polyA_signal 2908..2913
polyA_site 2927
/notes="putative"
/notes="putative"
BASE COUNT 837 a 581 c 664 g 845 t
ORIGIN
Query Match 30.7%; Score 628.2; DB 192; Length 2927;
Best Local Similarity 84.3%; Pred. No. 1.3e-149;
Matches 734; Conservative 0; Mismatches 128; Indels 9; Gaps 2;
QY 572 tgagttctggccttcagttatgatagatgcttagtttgaccttcctcagtgataaa 631
Db 1376 TGCTTTTCCCTTGTTAGTATGATGAATGGCTAGTGTTCAGCTTCCAGCTGTGATAA 1435
QY 632 ctttattttgcagaaacggccagaaagatctattatgtcggctattcagaggcac 691
Db 1436 CTTTATCTTACAGAAACGGCCAGAAAGGCTATTATGTGCGCTACTCACAGGGCAC 1495
QY 692 caccatgggtttattgatttccaccatgccagagctggtcagaaatacaaaatgta 751
Db 1496 CACCATGGGCTTATTGCATTTTCCACAATGCCAGAGCTAGCTACATAAAATCAAAATGTA 1555
QY 752 ttttgcttagcccatgacctgttaagcatgcaaaagcccgaggccaaattttt 811
Db 1556 TTTTGCCCTTAGCTCTATACCCACTGTAAATATGCAAGAGTCTCGTACCAAAATCTT 1615
QY 812 gttgctccagatgatgatcaaggagttgtttggcaaaagaaatttctgtatcacag 871
Db 1616 GCTGCTCCAGATATGATGATCAAGGTATTATTGGCCAGACAAGAGTTTATACCCAGAC 1675
QY 872 cagatttctcagacaacttctatttacccttctgtgcccaggtgatttctgacagattg 931
Db 1676 TAGATTTTTCAGACAGCTTTTATTATACCTTTGTGCGCAGATGATTCTTGACCAAAATCTG 1735
QY 932 tagtaatatcatgttacttctgggtgattcaacacaaacaataatgaacatgagccgagc 991
Db 1736 CAGCAACATCATCTTACTCTGGGAGGTTTACACAAACAATATGNAATGAGCAGAGC 1795
QY 992 aagtgtatatgctgccacactcttctggaacatctgtgcaaaatattctacactggag 1051
Db 1796 AAATGTATGTGCGCCATACACTGCTGGAACGCTGTGCAGAACATTTCTCCACTGGAG 1855
QY 1052 ccaggcagtaaatctgtgaactcgggcaatttgactgggggagtgagacaaataatct 1111
Db 1856 CCAGGCGGTGAATCTCGGGAACTCTGTCCTTTGACTGGGGAAGTGAGACCAAAATCA 1915
QY 1112 ggaataatgcaatcagccaaactcctgtaaggtacagagtcagagatgacggtccctac 1171
Db 1916 GGAAATATGTATACAGCCACTCTTATAGGTACAAAGTTTCAGATATGATGTGCCAAC 1975
QY 1172 agcaattgagacagggttcagactggtttcacaataccagaagacgtgaaaatgctgct 1231
Db 1976 AGCAATGTGACATGAGGTCAAGACTGGCTTTCAATCCAGATGATGTGAAACATTA 2035
QY 1232 ctctgaagtgacaaacctctaccataagaataattctcgaatgggctcagctgattt 1291
Db 2036 TTCTGAAGTACCAACCTCATCTACCAACAAGAACATTTCTGAGTGGGCTCATGTGATTT 2095
QY 1292 catctgggttttgatgctcctcaccgtatgtacaatgaataatcatcatctgatgagca 1351
Db 2096 CATCTGGGGCTGGATGCCCTCAGCGGTGTTTACAAATGAATTAATACATCTGATGAGC- 2154
```

```
QY 1352 ggaggagaccacaccttccaggagcggtgtgagccgtattgtgaagcatctgacactg 1411
Db 2155 --AGGAGCCCAACCTTCCCGCAGGAACTCGAGGTCAAATTTGGAAGCTCTCTGCAA--- 2209
QY 1412 acgatattgagacaacacctcctcctgaggatggg 1442
Db 2210 ---GTCTTAGAACAAACCCCATCGGAATGGG 2237
RESULT 2
BF135102 954 bp mRNA EST 24-OCT-2000
LOCUS 601780142F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4008131 5',
DEFINITION mRNA sequence.
ACCESSION BF135102
VERSION BF135102
KEYWORDS EST.
ORGANISM Mus musculus
SOURCE house mouse.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLN9242 row: c column: 12
High quality sequence stop: 695.
FEATURES
Location/Qualifiers
source 1..954
/organism="Mus musculus"
/strain="C25CH II"
/db_xref="taxon:10090"
/clone="IMAGE:4008131"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dr. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 280 a 222 c 222 g 230 t
ORIGIN
Query Match 24.7%; Score 505; DB 144; Length 954;
Best Local Similarity 85.3%; Pred. No. 3e-118;
Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;
QY 618 cctgagtgataaaactttatttgcagaaaaacggccagaaaaagatctattatgctcggc 677
Db 1 CCAGTGTGTGATAAATCTGATCTTACAGAAAACGGCCAGAAAAGGCTTATTATGTCGGC 60
QY 678 tattcacaggccaccaccatggctttattgctttttccaccatgccagagctggtcag 737
Db 61 TACTACAGGGCACCACCACCTGCTTATTGTCATTTTCCACATGCCAGAGCTAGCTCAT 120
QY 738 aaaaacaaatgtatttggcttttagcaccatagcactgttaagcagcaaaaaagcccc 797
Db 121 AAATCAAAATGTATTGCTTGTAGCTCTATAGCCACTGTAAATATGCAAGAGTCTCT 180
QY 798 ggaacaaatttttggctgcccagatgatcatcaaggagattgttggcaaaaaagaa 857
```

181	GGTACC	AAATTTCTGCTGCTGCAGATATGATGATCAAGGTATATTTGGCACAAGAG	240
858	tttctg	tacagaccagagattcttcagacaacttggttatttaccatttggccaggtgatt	917
241	TTTTTATAC	CAGACTAGATTTTTCAGGCAGCTTTTATTATTC-TTGTGGCCAGATGATT	299
918	cttgatc	gaatttgtagtaatacatgcttactctgggtgagttcaacacacaatatg	977
300	CTGCAC	AAATCTGCAGACACATCACTTACTCTGGAGGTTTAAACACAACAATATG	359
978	aacatg	agccgagcaagtgatatgctgccacacactcttctgctggaacatctgtgcaaaat	1037
360	AACATG	CAGCAGACAAATGTTATGCGCCCATACACCTGCTGGAGCGTCTGTGCAGAAC	419
1038	attctac	actggagccagcgagtgaaattcttggtgaactcgggcatttgactggggagat	1097
420	ATTCTCC	ACTTGGACGACGACGCCGTGAAATCTTGGGAACTTCGTGCCCTTTGACTGGGGAAT	479
1098	gagac	caaaaatctggaaaaatgcaatcagccaaactctctgaaggacagagtcagagat	1157
480	GAGAC	CAAAATCAGGAGAAATGTATACGCCAACCTCTATAGGTACAAAGTTCGAGAT	539
1158	atgac	ggtccctacagcaatgctggacagggaggtcaggactggttttccaaatccagaagac	1217
540	ATGATG	TCCCAACAGCAATGTGACTGGAGGTCAAGACTGG-TTTCAAATCCAGATGAT	598
1218	gtgaaa	atgctctctcctgaggtgaccaaactcatctaccataagaatatctctgaatg	1277
599	GTGA	AACTTACTTCTGAAGTAAACCAAACTCATCTPACCCACAGAACAATCTCTGAGGTG	658
1278	gctc	agtggaattcatctggggttttggatgctcctcaacg	1318
659	GGTC	ATGTGA-TTTCATCTGGGGGTGGATGGCCCTCAGCG	698

RESULT	3	
BF141567		
LOCUS		
DEFINITION		24-OCT-2000
		EST
		BF141567 1079 bp mRNA
		601788654F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5',
		mRNA sequence.
ACCSSION		BF141567
VERSION		BF141567.1
KEYWORDS		GI:10980607
SOURCE		house mouse.
ORGANISM		Mus musculus
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE		1 (bases 1 to 1079)
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/ .
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL		Unpublished (1999)
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgaps@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9263 row: 1 column: 14 High quality sequence stop: 663. Location/Qualifiers 1. .1079 /organism="Mus musculus" /strain="CZECH II" /db_xref="taxon:10090" /clone="IMAGE:4016413" /clone_lib="NCI_CGAP_Lu30" /tissue_type="tumor, metastatic to mammary" /lab_host="DH10B" /note="Organ: lung; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: SalI; transgenic model WNT-1, expression driven by
FEATURES		
source		


```

Db 850 AAAACTTAATGTGAGTCGGCTTTGGTGTGTATCTAGGCGCATAAATCCACGAGACATCTA 909
Qy 1030 tgcataattattacactgagccagcgagtgaaattcttggtgaactcgcggcatttgact 1089
Db 910 CTCAGAGACTTTTCCACTGGCCACAGCTTGCTAAATCTGGGAAGCTTCAAGCCTTAAC 969
Qy 1090 ggggagtgagacaaataatgcgaataatcaatcagcccaactccctgtaaggtagacag 1149
Db 970 GGGGAAGTCCATTCACAGAACATGTTACTACTACAATCAGAAACGCCCTCCCTACTATGATG 1029
Qy 1150 tcagagatatgacggtccctcagacaaatgtgacagaggtcagagctgggtttcacaatc 1209
Db 1030 TGTGAGCCATCAGCTGCAATTCAGTGTGGAAGCGGTGCGCATGACATCTGGCTGATC 1089
Qy 1210 cagaagacgtgaaatgctgtctctcaggtgagacacacctcatctaccataagaatattc 1269
Db 1090 CCCAAGATGTCGCAATGCTGCTCCCAACATCCCAACCTTCTGTACCATGAAGGATTC 1149
Qy 1270 ctgaatggggtcacgtgattctcctgggttttggtgctcctcaccgtatgtacaatg 1329
Db 1150 TTCCCTACATCAGCTGCAATTCAGTGTGGAAGCGGTGCGCATGACATCTGGCTGATG 1209
Qy 1330 aaatcatccatctgatg 1346
Db 1210 AGATAGTTACCATGATG 1226

RESULT 7
LOCUS AK009300 1349 bp mRNA HTC 08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310011G18, full insert sequence.
ACCESSION AK009300
VERSION AK009300.1 GI:12844011
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:2310011G18.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Tamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Washiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Kashiiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1349)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Horii,F.,

```

Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGACGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 5.0 and subtraction to Rot - 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATCTCGAGTAAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES
source
Location/Qualifiers
1..1349
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1895751"
/db_xref="MGD:MGI:1914967"
/clone="2310011G18"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/note="putative"
/codon_start=1
/protein_id="BAB26201.1"
/db_xref="GI:12844012"
/translation="MWLLVTSVLSAFGGAGHGLFGKLPKNPEANNVNQMYWGYSPSEYEVVTEGYILGVYRIPYGRKNSENIGRPVAYLQHLIASATNWTILPNNLSLA FILADAGYDVLGNRSRKNVYSPDSVEFAFSDEMARYDLDPATIDFIVQTK GQRKHVHGSGQTIGFIAFTNPALAKIKREYALAPVATVRYKPKSLIPK FLKIVFGNKMFPNHYLDQFLGTEVCSRELLDLCNALFICGDFDKNNLVSREGV YLGNPAGTSTQDLFWAQLAKSKLQYNGWSPLQNLMLHYNQTPPYDVSATVPI AEWNGGHDIADPDQVAHLPLKPLNLLHKLPLNLLDFIWMADPAQVEYNEIVTM AED"

CDS
395 a 290 c 284 g 380 t

BASE COUNT
ORIGIN
Query Match 22.5%; Score 460.2; DB 192; Length 1349;
Best Local Similarity 63.7%; Pred. No. 9.3e-107;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

Qy 250 accagaagcattcatgaattagtgaaatcatccacatcaatcctcctgtgagg 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132 ACCCTGAAGCAACATGAATGTTACTAGATGATCACTTACTGGGATATCCAGTGAGG 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 310 aatgaagtcgcaactgaatgggtatctcttctgttaacaggattcctcgagcc 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 AATATGAAGTGTGTACTAGAGATGCTACATTTCTGGGGGTCTATAGAAATCTTATGGGA 251

Qy	370	tagtgcacctaaagaacagagtctccaggcctgtggtttactgcacgatggcgcttagtg	429
Db	252	AGAAAAATTTCTGAGAATATTCGCAAGACAGACCTTGGGCATATTTCGACCATGGTTGTGATG	311
Qy	430	gagggtctagcaactggatttccaacctgccacaacatagcctgggttcattctgcagcag	489
Db	312	CATCAGCCACNAACTGGATTACAATCTGCCAACACAGCGCTGGCCTTCAITTCATCTAGCAG	371
Qy	490	atgctgggttttgacgtgtggatgggggaacacagcagggggaacgcctggtctcgaatacacca	549
Db	372	ATGCTGGCTATGATGTGCGCTGGGAACACGTCGAGGGAATACATGTCCTCCGGAAAATG	431
Qy	550	agacactctccatagaccaaagtgggtctgtgggtcttcagttatgatgatgggtcaggt	609
Db	432	TATACTATTTCACCAGCACTCAGATTGAATTTGGGCTTTTCAGCTTTTGATGAATGGCTAAAT	491
Qy	610	ttaactctctcagtgataaacttattttgcagaacggccaggaataaatctatt	669
Db	492	ATGACCTTCAGCCACCATTAGACTTCATTGTACAGAAACTGGACAGAGAAGATACACT	551
Qy	670	atdctgggtattccagggcccaccaccaatggggtttattgcattttccaccactgcacagc	729
Db	552	ATGTTGGTCACCTCTCAGGGCACCACTATCGGTTTTATTGGCCTTTTCACCAATCCTGCTC	611
Qy	730	tggctcagaataacaaaatgtattttgtcttagcaccatagccactgtttaagcatgcaa	789
Db	612	TGGCTTAATAAAATCAAGAGGGTTTTATGATTAGCTCCAGTTGCTACTGTGTAAGTATACAG	671
Qy	790	aasgccccggaccaaatttttgtgctgcacagatatgatcaagggtattgttgcca	849
Db	672	AAAGTCCCITTAANAAGATTTCACITTATCTCTAAGTTTCTCAAGGTGATATTGGTA	731
Qy	850	aaaaaattcttgtatcagacagagatttcttcagacaacttggttatccotttgtygcc	909
Db	732	ACAAAATGTTTCATGCCCCACAACTACTTAGATFCAAATTTCTTGGTACGGAACTGTGCTCAC	791
Qy	910	aggtgatctcttgatcagatttgttagtaatacatactgtactctcctgggtggatccaacca	969
Db	792	GGAGCTGCTAGATCTCTCTCGACNACCGCTTATTCATCTTCTGTGGATTTGACAAGA	851
Qy	970	acaaatgaacatgagccgagcaagtgtatatgtgcccaactcttctgtggaacatctg	1029
Db	852	AAAACTTAAATGFGAGTCGCTTTTGGTGTATCTTAGGGCATAATCCACGAGAACATCTA	911
Qy	1030	tgaasaatattctaactggagccaggcaagtgaaattctcgttgaaactocggscatttgact	1089
Db	912	CTCAAGACCTTTTCCACTGGGCACAGCTTGCTTAAATCTGGGAAGCTTCAAGCCTATAACT	971
Qy	1090	gggggagtgagacaaaatctggaanaatgcaatcaagccaaactcctgtgaaggtcacagag	1149
Db	972	GGGAGTCCANTACAGACATGTTTACACTACATCAGAAACCGCTCCCTACTATGATGATG	1031
Qy	1150	tcagagatatgacggtccctacagcaatgtgacaggaggtcagactggotttcaaatc	1209
Db	1032	TGTACGCCATGACCGTGCCAAATTCAGTGTGGAACGGTGGCCATGACATCCTGGCTGATC	1091
Qy	1210	cagaagacgtgaaaatgctcgtctctctgtgagtgaccaaacctcatctacataagaatttc	1269
Db	1092	CCCAAGATGTGCGAATGTGCTTCCCAACTCCCCAACCTTCTGTACCATAAGGAGATTC	1151
Qy	1270	ctgaatgggctcagtggaatttcattctgggttttgatgctccctcacogtatgacaatg	1329
Db	1152	TTCCCTACAAATCAGCTGGACTTCATCTGGCGCATGGATGGCCCTCAGAGGTTTACAATG	1211
Qy	1330	aaatcattccatctgctg	1346
Db	1212	AGATAGTTACCATGATG	1228

library, clone:2310069p19, full insert sequence.

ACCESSION AK010139

VERSION AK010139.1 GI:12845371

KEYWORDS CAF trapper.

SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library

clone:2310069p19.

Mus musculus

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (sites)

REFERENCE Carninci,P. and Hayashizaki,Y.

AUTHORS High-efficiency full-length cDNA cloning

TITLE Methods Enzymol. 303, 19-44 (1999)

REFERENCE 2 (sites)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

REFERENCE 3 (sites)

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,Y., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

REFERENCE 4 (sites)

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and PANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5 (bases 1 to 1350)

AUTHORS Adachi,T., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Koizumi,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishit,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Soqabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

RESULT	8
LOCUS	AK010139
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched
	1350 bp mRNA HTC 08-FEB-2001

Host: SOLR.

FEATURES
source

Location/Qualifiers
1. .1350
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1898377"
/db_xref="MGD:MGI:1914967"
/clone="2310069P19"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53. .1240
/note="putative"
/codon_start=1
/protein_id="BAB26725.1"
/db_xref="GI:12845372"
/translation="MWLLVTSVLSAFGGHGLFKLPKPNPEANMNVSONITYWGY
SEETVEDGILGVIRIPYKKNSENIGRPVAYLOHGLIASATNINLNPNSLA
FILADAGYDVLGNRSNTWSRKNVYSPDSVEFAFSDEMAKYDLPATIDFVQKT
GOEKIHVYGSQGTIGFIAFSTNPALAKIKRFLAPALATVATKYTESPKKISLIPK
FLKAVIFGNKMFPHNYLDQFLGTEVCSRELLDLCSNALFIFCGPKKNLVNRFEDV
YLGNPAGTSTQDLFHNQALAKSKLOAYNWSPLQNLNHNOKTPPYDVSANTVPI
AWNGGHDLADPDQDVAMLLPKLPNLLYHKEILPYNHLDFIWAMDAPQEVYNEIVTM
AED"

CDS

BASE COUNT 398 a 289 c 283 g 380 t

ORIGIN

Query Match 22.5%; Score 460.2; DB 192; Length 1350;
Best Local Similarity 63.7%; Pred. No. 9.3e-107;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 250 acccagaagcattcatgaattagtgaaatcatcaatcaatcaatcaatccctctgagg 309
DB 132 ACCCTGAAGCAACATGAATGTAGTCAGATGAATCACTTCTGGGGATATCCAGTGAGG 191

QY 310 aatatgaagtcgaactgaagtggttatatctcttcttgaacaggatctctcgagcc 369
DB 192 ATATGAAGTTCTTACTGAAGATGCTACATCTCTGGGGCTATAGAATTCCTTATGGGA 251

QY 370 tagtgcaacctgaagaacaggtccaggcctgtgtgttaactgcagcatggcctagttg 429
DB 252 AGAAAAATCTGAGAAATTCGCAAGAGACCTGTGGCATATTTCAGCAGATGTTTGATG 311

QY 430 gagggtcagcaactggtttccaccctgcccaacaatagctggcttcattctcggcag 489
DB 312 CATCAGCCCAACACTGGATTACAAATCTGCCAACAACAGCCTGGCCTTCATCTAGCAG 371

QY 490 atgctgggtttgacgttgatggggaacagcagggggaacccctggctcgaacacaca 549
DB 372 ATGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGGAATACATGGTCCCGGAAATG 431

QY 550 agacactctccatagaccagaatgagttctggccttcagttatgatgagatggcaggt 609
DB 432 TATACTTACCACAGACTCAGTTGAATCTGGGCTTTCAGCTTTTCATGAATGGCTAAT 491

QY 610 ttgacctctcagtgataaactttatttgcagaaacggcaggagaaagactctatt 669
DB 492 ATGACCTTCCACCCACCATAGACTTCATTTGTACAGAAAACCTGGACAGAGATACACT 551

QY 670 atgctggctattcagggcaccaccatggcctttattgatttccaccatgccagagc 729
DB 552 ATGTTGGTCACCTCTCAGGCGCACCATATGCGGTTTATTGGCTTTCTTACCAATCTCGTC 611

QY 730 tggctcagaaaaatcaaaatgtatttgccttagcaccatagccactgtagcatgca 789
DB 612 TGGCTAAAAAATCAAGAGGTTTTATGCATTAGCTCCCAATTCCTACTGTGAATACAG 671

QY 790 aaagcccccgggacccaaatatttctgtgcagatatgatcaagggattgttgca 849
DB 672 AAAGTCCCTTTAAAAAGATTTTCACTATTTCTTAAAGTTTCTTCTCAAGGTGATATTGGTA 731

QY 850 aaaaagaatttctgtatcagaccagatttctcagacaacttcttattacctttgtggcc 909
DB 732 ACAAAATGTTTCATGCCCCCACAACTACTTAGATCAATTTCTTGGTACGGAAGTGTCTCAC 791

QY 910 aggtgattcttgatcagatttgtagtaataataatcttactcttgggttgattcaacaca 969
DB 792 GGGAGCTGCTAGATCTTCTCTGCACCAACGCTTTTATTCATCTTCTGTGGATTGACAAGA 851

QY 970 acaatatgaacatgagccgagcaagtgtatatgtctgctgcccacactcttctggaacatctg 1029
DB 852 AAAACTTAATGTGAGTCGCTTTGATGTGATCTAGGCGATATTCAGCAGGAGAACATCTA 911

QY 1030 tgcataatattctacacdgagccaggcagtgaaattcttggtgaactccgggacttgact 1089
DB 912 CTCAGACCTTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGCCTATAACT 971

QY 1090 gggggagtgagacaaaataatctggaataatgcaatcagcccaactcctgtaaggtacagag 1149
DB 972 GGGGAAGTCCATTACAGAACATGTTTACACTACAAATCAGAAAAACGCTCCCTACTATGATG 1031

QY 1150 tcagagatatgacggctccctcagcaatgtgacaggaaggtcagagctggtcttcaaatc 1209
DB 1032 TGTACCCATGACCTGCGCATTTGAGTGTGGAAAGGTGGCCATGACATCTCGGCTGATC 1091

QY 1210 cagagacgtgaaaatgctctctctgaggtgaccaacactcattaccataagaatattc 1269
DB 1092 CCCAGATGTGCAATGCTGCTTCCCAACTCCCAACCTTTGTACCATAGGAGATTC 1151

QY 1270 ctgaatgggctcaactgagattctctctggtttggtgctcctcaccctatgtacaatcg 1329
DB 1152 TTCCCTACAAATCACCTGGACTTCTATCTGGCGATGATGCGCTCAAGAGGTTTACAATG 1211

QY 1330 aaatcatccatctctgatg 1346
DB 1212 ACATAGTTACCATGATG 1228

RESULT 9
AK010124

LOCUS AK010124 1340 bp mRNA HTC 08-FEB-2001

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069D09, full insert sequence.

ACCESSION AK010124

VERSION AK010124.1 GI:12845346

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain: C57BL/6J) adult male tongue cDNA to mRNA, clone: 2310069D09.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (sites)

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Methods Enzymol. 303, 19-44 (1999)

REFERENCE 2 (sites)

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

REFERENCE 3 (sites)

AUTHORS Shibata, K., Itoh, M., Alizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, N., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

Db 1146 TTCCTACAATCACCTGGACITCATCTCGCGCATGGATCGCTCAGGAGGTTTACAATG 1205

QY 1330 aaatcatcatcatgtatg 1346

Db 1206 AGATAGTTACCATGATG 1222

RESULT 10

AK009479 1343 bp mRNA HTC 08-FEB-2001

LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022K08, full insert sequence.

ACCESSION AK009479

VERSION AK009479.1 GI:12844303

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:2310022K08.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (sites) Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Methods Enzymol. 303, 19-44 (1999)

AUTHORS 2 (sites)

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

AUTHORS 3 (sites)

Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

AUTHORS 4 (sites)

The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

AUTHORS 5 (bases 1 to 1343)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Iwata, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SstI.

FEATURES

Location/Qualifiers

1. 1343

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="MGD:MGI:1896992"

/db_xref="MGD:MGI:1914967"

/clone="2310022K08"

/sex="male"

/tissue_type="tongue"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

50. 1237

/note="putative"

/codon_start=1

/protein_id="BAB26316.1"

/db_xref="GI:12844304"

/translation="MWLLVTSVLSATGAGHGLFGKLPKPNPEANMNVNQMTWYGYG SEYEVEDGYILGVIRPYGKNSNENIGRPVAYLOHGLIATNWTNLPNNLSLA FILADAGYDVLGNSRGTNRKNVYSPDSFWMFADYKATIDFIVQKT GOEKIHYGVHSGQTIGFIATFSTNPALAKIKRYALAPVATVRYTESPFKKISLIPK FLKVIYGNKMFPHNYLDQFLGTEVCSRELLDLCSNALIFGDKKLNLSRFDV YLGNHPAGTQDILFHWALQASGLQAYNWSPLONMLHYNOKTPPYDYDSAMTPVI AWVNGGHDLIADPDQVAMLLPKLPNLYLHKELPYNHLDFIWMADAPQEVINEITVM AED"

BASE COUNT 394 a 289 c 282 g 378 t

ORIGIN

Query Match 22.4%; Score 458.6; DB 192; Length 1343;

Best Local Similarity 63.6%; Pred. NO. 2.4e-106;

Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 250 accagaagcattcatgaattagtgaaatcatcaacatcaaggctatccctgtgag 309

Db 129 ACCGTGAGCAACATGAATGTTAGTCAGATGATACCTTCTGGGATATCCAGTGAGG 188

QY 310 aatagaagtcgaactgaagatgggtatatactcttctgttaacaggattctctgagcc 369

Db 189 AATATGAAGTTGTTACTGAAGATGCTACATTTCTGGGGCTATAGAAATCTTATGGGA 248

QY 370 tagtgcacacttaagaagcagggtccaggctgtgtgttactgagcagtgagctagtgtg 429

Db 249 AGAAATAATTCAGAAATATCGCAAGACGACCTGTGGCATATTTGCAGCATGTTTGAATG 308

QY 430 gaggtgtgacactggatttccaaaccctcccaacaaatagcctgggtctcattctgagc 489

Db 309 CATCAGCCACAACTGGATTAACAATCTGCCAACACACAGCTGGCTTCATCTTACGAC 368

QY 490 atgtcgtgtttgacgtgtggtggggaacagcagcggggaacgcctgggtctcgaacaca 549

Db 369 ATGTGCTATGATGCTGTGGCTGGGAACAGTCGAGGGAATACATGTGCTCCCGGAAATG 428

QY 550 agaacactccatagacaaagatgagttctgggtcttcagttatgatgagatggctaggt 609

Db 429 TATATCTATTCACGACACTCAGTTGAATTCCTGGGCTTTTCAGCTTTGTGTAATGGCTAAAT 488

QY 610 ttgacctctcgtcagtgataaactttatttgcagaaaaacggccaggaagaaactatt 669

Db 489 ATGACCTCCAGCCACCATGACTTCATCTGTACAGAAAATCGAACAGAGAGATACACT 548

Qy	670	atgtcggtctattccaggccaccacatgggctttattgcatctttccaccatgccagagc	729
Db	549	ATGTGTGTCTACTTCACGGGACCACTACTCGGTTTTATTGCTTTTCTACCAATCTGCTC	608
Qy	730	tggctcgaataacaaatgtatttcttagcaccatagccacgtttaagatgcaa	789
Db	609	TGGCTAAATAATCAAGAGGTTTTATGATTAGCTCCAGTTGCTACGTGAAGTATACAG	668
Qy	790	aaagcccccggaccacaaatttggctgcccagatatgatcatcaagsgattgttgcca	849
Db	669	AAAGTCCCTTTAAAAAGATTTCATTATCTCTAAGTTCTTCTCAAGGTGATATTGGTA	728
Qy	850	aaaaagaattctctgatacagaccagatctctcagacaaacttgatttaccotttggtcc	909
Db	729	ACAAAATGTTTCATGCCGCCACAACACTACCTAGATCAATTTCTTGGTACGGAAGTGTCTAC	788
Qy	910	agtgattcttgaccagattgttagtaatacatgttacttctgggtgagattcaaccca	969
Db	789	GGGAGCTGCTAGATCTCTCTGCAGCAACGGTTTATTCAATCTCTGTGGATTGGACAGA	848
Qy	970	acaatatgaacatgagccgagcaagtgtatatgtgcccacactcttgctggaacatctg	1029
Db	849	AAAATTTAAATGTGAGTCGCTTTGATGTGTATCTAGGCATAATCCAGCAGAACATCTA	908
Qy	1030	tgcataattctacatgagccagcagtgaaattctgtgaaactccggcatttgact	1089
Db	909	CTCAAGACCTTTTCCACTGGCACAGCTGTGTAATCTGGGAAGCTTCAAGCCTATAACT	968
Qy	1090	ggggagctgagaccacaaatctggaataatgcaatcagccaactccgtgaagctacagag	1149
Db	969	GGGAAGTCCATTACAGACATGTTTACACTACAATCAGAAACGCCCTCCCTACTATGATG	1028
Qy	1150	tcagagatatgacggtccctacagcaatgtggacagaggtcagagactgggtttcaaatc	1209
Db	1029	TGTCAGCATGACCGTGCCAAATGTCAGTGTGGAACGGTGGCCATGACATCTCGGCTGATC	1088
Qy	1210	cagaagacgtgaaatgctctctctgagtgaccaacctcatcaccataagaatttc	1269
Db	1089	CCCAAGATGTCGAATGCTGCTCCCAAACTCCCAACCTTCTGTACATAAGCAGATTC	1148
Qy	1270	ctgaatgggtcagctgagattctatctgggtttggatgctctccacgtagtacaatg	1329
Db	1149	TTCCCTACAATCACTCGACTTCATCTGGGCGATGGATGGCCCTCAAGAGGTTTACAATG	1208
Qy	1330	aaatcatccatctgatg	1346
Db	1209	AGATAGTTACCATGATG	1225
RESULT 11			
AK009413			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

FILADAGYDVLGNRGNWTSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKT
 GOEKIHVHSGQGTIGFIAPFNPALAKKIKRYPALAPVATVKYTESPKKISLIPK
 FLKIFGNKHEPHNYLDQFLATEVCSRELLDLCSNLFIFCGFKNLNMYSDREY
 YLGNHPAGTSTQDLFWAQLAKSGKLQVNWGSPLONLHYNOKTPPYDVSAMTVP
 AVWNGHDILADPODVANLPLKPLNLLYHKEILPYNHLDFIWMADAPQEVNEYVTMM
 AED

BASE COUNT 395 a 288 c 283 g 379 t

Query Match 22.4%; Score 458.6; DB 192; Length 1345;
 Best Local Similarity 63.68; Pred. No. 2.4e-106;
 Matches 598; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 250 acccagaagcattcatgaattagtgaaatcccaacatcaaggctcccttgtagg 309
 DB 131 ACCCTGAAGCAACATGAATGTAGTCAGATGATACTTACTGGGGATATCCAAGTGAGG 190
 QY 310 aatataagtcgcaactgaagtggtatatacctcttctgtaaacaggattcctcgagcc 369
 DB 191 AATATGAAGTGTGTAACTGAAGTCTACATCTCGGGGTCTATAGAATCTCTTATGGGA 250
 QY 370 tagtgcaacctaaagaagcaggttccagcctgtgtgttactgcagcatggcctagtg 429
 DB 251 AGAAAAATCTCGAGAAATATCGCAGAGACCTGTGGCATATTTGCAGCATGGTTTCATG 310
 QY 430 ggggtctagcaactggtattcccaacctgcccacaacatagcctgggtctattctggcag 489
 DB 311 CATCAGCCACAACTGGATTACAAATCTGCCAACACAGCGCTGCCCTTCTATCTAGCAG 370
 QY 490 atgctgggttgacgttggtatgggaagcagcaggggaagcctggtctcgaaaaacaca 549
 DB 371 ATGCTGGGTATGATGTGTGGCTGGGGAACAGTCGAGGAATAACATGTCCTGCTTCTG 430
 QY 550 agacactctccatagaccacgaatggtctgggtcttgcagaaacggttcagtgagtgag 609
 DB 431 TATACTATTCACACACTCAGTTGAATCTGGGCTTTCAGCTTTCAGTAATGCTTAAT 490
 QY 610 ttgacctctcgcagtgataaacttatttgcagaaacggcagggaagaagattctatt 669
 DB 491 ATGACCTTCCAGCCACCATAGACTTCATTGTACAGAAAACCTGGCAAGAGATACACT 550
 QY 670 atgctgggtcttaccagggccaccacattggccttattgcatcttttccaccatgccagagc 729
 DB 551 ATGTTGGTCACTCTCAGGCGCACCATATCGGTTTATTGGCTTTCTTACCAATCCTGCTC 610
 QY 730 tggctcagaaatcaaaatgtatttggctttagcaccatagcactgttaagcatgcaa 789
 DB 611 TGGCTAAAAAATCAAGAGGTTTATGATATGATTCAGTTCAGTGTCTACTGTGAATACAG 670
 QY 790 aaagcccccgggaccacaaattttgtgtgcagatatgatcatcaagggttgggtgca 849
 DB 671 AAAGTCCCTTTAAAAAGATTTCTACTTATTCCTAAGTTTCTTCTCAAGGTGATATTGCTA 730
 QY 850 aaaaagaattctgtatcagaccagattcttcagacaactgttattacatttgggccc 909
 DB 731 ACAAAATGTTATGCCCCACACTACTTAGATCAATTTCTGGTACGGAAGTGTGCTCAC 790
 QY 910 aggtgatctcttgatcagattgttagtaataatcatgttactctgggtggtattcaacacca 969
 DB 791 GGGAGCTGTAGATTCTCTGCACACAGCGTTTATTCATCTCTCTGTGGATTGTACAGA 850
 QY 970 acaatataaactatgagccgagcaagtgatatgtgctgccacactctgtcggaacattcg 1029
 DB 851 AAAAATTAATGTGAGTCGCTTGTATGTATCTAGTGGCAATATCCAGCAGGAACATCTA 910
 QY 1030 tgcacaaatattacactggagcagcagtgaaattcttggtgaactccggcatttgact 1089
 DB 911 CTCACAGACTTTTCCACTGGGCACAGCTGTGTAATCTGGGAAGCTTCAAGCCTATAACT 970
 QY 1090 gggggagtgagaccacaaatctggaaaatgcaatcagcgaactcctgtaagggtcacag 1149
 DB 971 GGGGAAGTCCATTACAGAACATGTTTACTACTACATACAAATCAGAAAACGCCCTCTACTATGATG 1030

QY 1150 tcagagatatgacggtcctccacagcaatgttgacagagggtcagagctgtttcaaatc 1209
 DB 1031 TGTGAGCCATGACCGCTGCAATTTGAGTGTGGAACGGTGGCCATGACATCTCTGGTGTATC 1090
 QY 1210 cagaagacgtgaaatgtctctctgaggtgagcaaacctcattcaccataagaattatc 1269
 DB 1091 CCCAAGATGTGCAATGCTGCTTCCCAAACTCCCAACCTCTTGTACCATAGGAGATTC 1150
 QY 1270 ctgaatggggtcagctggtgattcattctgggtgttgagctcctcaccgtatgtacaatg 1329
 DB 1151 TTCCCTACAACTCAGCTGACTCTCATCTGGCGATGGATGGCGCTCAAGAGGTTTACAATG 1210
 QY 1330 aaatcatccattctgatg 1346
 DB 1211 AGATAGTTACCATGATG 1227

RESULT 12
 AK009473 1345 bp mRNA HTC 08-FEB-2001
 LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 DEFINITION library, clone:2310022H17, full insert sequence.
 ACCESSION AK009473
 VERSION AK009473.1 GI:12844293
 KEYWORDS CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
 clone:2310022H17.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (sites)
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Methods Enzymol. 303, 19-44 (1999)
 REFERENCE 2 (sites)
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 REFERENCE 3 (sites)
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
 Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
 Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
 Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipillar sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 REFERENCE 4 (sites)
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5 (bases 1 to 1345)
 AUTHORS Atachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
 Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
 Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
 Inotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
 Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
 Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
 Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
 Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
 Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 TITLE Direct Submission

[illegible]

		TITLE	JOURNAL	COMMENT
		<p>Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)</p>		
		<p>Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.</p>		
		<p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACAGATCCAGACTCGTCTTTTCTTTTTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTAAATTAAATTAATGCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.</p>		
FEATURES	SOURCE	Location/Qualifiers		
		1..1346		
		/organism="Mus musculus"		
		/strain="C57BL/6J"		
		/db_xref="taxon:10090"		
		/db_xref="MGD:MGI:1895759"		
		/db_xref="MGD:MGI:1914967"		
		/clone="2310028109"		
		/sex="male"		
		/tissue-type="tongue"		
		/clone_id="RIKEN full-length enriched mouse cDNA library"		
		/dev_stage="adult"		
	CDS	53..1240		
		/note="putative"		
		/codon_start=1		
		/protein_id="BAB26338.1"		
		/db_xref="GI:12844370"		
		/translation="IMLLLVTSVLSAFGAHGLFGKLGPKNPKNPEANMNVSONITWGPVSEEEAVTEDEGYILGVYRIPGKKNNNSINGRPAVAYSPHEMKAYIDLPIDIVQGTFLADAEVYDWLIGSRNNTSSRRKVVSSPSVPAMASPOBMAKYDLPIDIVQGTGOEKTHVHGSGCTTGTFIAFSTPDAIAKKIRYALAPVAATVTSPSKSLIRKFLVNGKMFPMFNHYLDLDFLGTEVSRLELDLCSNAIFEGCFKKNUNASRFQVYLGNPAGTSTODLFEMAOAKSGKLDAVMWSPLQNMHLHNOKTPRYDVSAATVPIAVNWGHDIILADPDQVAMLLPKLPNLNLYHKELIPLYNHDLFIWMADAQEVYNEIYTMWAED"		
BASE COUNT	395 a	289 c	283 g	379 t
ORIGIN				
Query Match	22.4%	Score 458.6;	DB 192;	Length 1346;
Best Local Similarity	63.6%	Pred. No. 2.4e-106;		
Matches 698;	Conservative	0;	Mismatches 399;	Indels 0;
Gaps	0;			
Db	132 ACCCTGAAGCAACATGATGTACTCAGATGATACTACTCTGGGAGATATCCAAAGTAGG	191		
Yy	310 aaatgaagtgcgaactaaagaatggtaatactccttcgtttaacaggagattcccgagacc	369		

Db 192 AATATGAAGTCTTACTAGAGATGGCTACATCTCTGGGGTCTATAGAAATTCCTTATGGGA 251
QY 370 tagtgcaacctaaagacaggttccaggcctgtgtgttactgtcagcatggcctagtgtg 429
Db 252 AGAAAAATCTGAGATAATCGCAAGACAGCTGTGGCATATTTGTCAGCATGGTTGATTG 311
QY 430 gagggtcgtacgaactggattttccaccctgcaccaacatagcctgggttcattctcggcag 489
Db 312 CATCAGCCACAAATCGGATTCAGAAATCGCCAAACAGACGCTGGCCCTTCATTTAGCAG 371
QY 490 atgtcgtgtttgacgtgtggaagggaacagcagggaacccctgggtctcgaataacaca 549
Db 372 ATGCTGGCTATGATGTGTGGCTGGGAGACAGTCGAGGGGATACATGGTCCCGAAATG 431
QY 550 agacactctccatagacaaagatgagttcgtggcttccagtttatgatgatgagtggt 609
Db 432 TATACATTTCCACAGACTCAGTTGAATCTCTGGCTTTTTCAGCTTTTATGAAATGGCTAAT 491
QY 610 ttgaccttctcgtcagtgataaactttatttgcagaaacagggccagaaagatctatt 669
Db 492 ATGACCTTCCAGCCACCATAGACTTCATTTGTACAGAAACTGGCAAGAGAGATACACT 551
QY 670 atgtcggctattcacaggccaccacatgggctttattgcatcttccaccatgccagagc 729
Db 552 ATGTTGGTCACTCTCAGGGCACCACATATCGGTTTATTGGCTTTCTTACCATCTGCTC 611
QY 730 tggctcagaataatcaaatgtatttctgttagcaccatagccactgttaagcatgcaa 789
Db 612 TGGCTAAAAAATCAAGAGTGTATGATGATAGCTCCAGTCTGCTACTGTGAAGTATACAG 671
QY 790 aaagcccgaggacaaattttgtgtgcgcagatatgatgatacgaaggatgtttgcca 849
Db 672 AAAGTCCCTTTAAAAAGATTTTCCATTTATCTTAAAGTTTCTTCAAGGTGATTTGGTGA 731
QY 850 aaaaagaattctgtatcagaccagacttctcagacaactgttatttaccctttgtggcc 909
Db 732 ACATAATGTTATGTCCTCCACACTACTTAGATCAATTTCTGGTAGGAGTGTCTCATC 791
QY 910 aggtgattcttgatcagattgttagtaatatcatgttactcttctgggtggttcaacaca 969
Db 792 GGGAGCTGTAGATCTCTCTCAGACAGCGTTTATTTCATCTCTGTGGATTGACAAGA 851
QY 970 acaatatgaactagccagcagcaagtgtatgtatgtcgcacacactcttctggaacatctg 1029
Db 852 AAAACTTAAATGTGATCGCTTTGATGTGTATCTAGGGCATATCCAGCAGGAGCATCTA 911
QY 1030 tcaaaaattctcactgagccagcagtgaaattctggtgaactccgggcatttgact 1089
Db 912 CTCAGACCTTTTCCACTGGGCACAGCTTGTCTAAATCTGGGAGCTTCAAGCCTATACT 971
QY 1090 ggggagtgagaccaaaactcggaaaaatgcgaatcagccaactcctgttaagggtacagag 1149
Db 972 GGGGAAGTCCATTACAGAACATGTTACACTACAAATCAGAAAAAGCGCTCCCTACTATGATG 1031
QY 1150 tcagagatatgaggtccctcagacaagtgtgacagaggaggtcagagactggttcaaatc 1209
Db 1032 TGTCAAGCCTTACCGTGGCAATTTGCAATGTGGAACGGTGGCCATGACATCTCTGGTGTG 1091
QY 1210 cagaagacgtgaaatgctcctctcgtgagtgaccaaactcctcattaccataagaatttc 1269
Db 1092 CCCAGATGTCGAATGCTGCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC 1151
QY 1270 ctgaatgggctcagctggatttcattcctgggtttggtgctcctcaccgtatgtacaatg 1329
Db 1152 TTCCCTACAAATCACCTGGACTTCATCTGGCGCATGGATCGGCTCAAGAGGTTTACAATG 1211
QY 1330 aaatcctccatctgatg 1346
Db 1212 AGATAGTTACCATGATG 1228

RESULT 15
AK009729

LOCUS AK009729 1346 bp mRNA HTC 08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310040L03, full insert sequence.
ACCESSION AK009729
VERSION AK009729.1 GI:12844702
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:2310040L03.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE 3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Ozawa,Y., Izawa,M., Ohara,E., Wakahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE 4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1346)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Sakai,K., Sano,H., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'

GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

source
1. 1346
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1898390"
/db_xref="MGD:MGI:1914967"
/clone="2310040L03"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53. 1240
/note="putative"
/codon_start=1
/protein_id="BAB26466.1"
/db_xref="GI:12844703"
/translation="MWLLVTSVLSAFSGAHGFLGKLPKNPEANMNVSMQITYWGYP
SEEVVTEDEYILGVYRIPYKNSENIGKRPVAYLQHLIASATNWTNLNPNNSLA
FILADAGVDVILGNSRGNTWSRKNVYSPDSVEFWAFSPDEMAKYDLPIDIVOKT
GOEKHYVHSGQTTIGFIASFNPALAKIKRYPALAPVATVYKYPSPFKKSLIPK
FLKVIETGNMFPHNYLDQFLGTEVCSRELLDLCSNALFICFGPKNKLNVSRFDV
YLGHNPAFTSQDLFWAQLAKSGKLOAYNMGSPLOMLHYNOKTPPYVDVSAMTVPI
AVWNGHDIILADPOVAMLLPKLPNLLYHREILPYNLHLPFIWAMDAPQEVYNEIVTMM
AED"

CDS

BASE COUNT 395 a 288 c 283 g 380 t
ORIGIN
Query Match 22.4%; Score 458.6; DB 192; Length 1346;
Best Local Similarity 63.6%; Pred. No. 2.4e-106;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;
QY 250 acccgaagcattcatgaattatgaatgaatcccaatcaatcaagcctccctgtgagg 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 132 ACCCTGAAGCAACATGAATGTTAGTCAGATGATAACTTACTGGGGATATCCAAGTGAGG 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 aatatgaatgcgaactgaatgggtatattccttctgttaacaggaggtccctgagccc 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 192 AATATGAAGTTGTTACTGAAGATGCTACATCTCTGGGGTCTATAGAAATCTCTATGGGA 251
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 370 tagtgcacctaagaagacaggttccaggcctgtgtgttactgcagcagatggcctagttg 429
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 252 AGAAAAATTCAGAAATATCGCAAGAGACCTGTGGCATATTGGCAGCATGGTTTGATG 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 430 gagggtgtagcaactggatttccaccctgcccaacaatagcctgggtctcattctctggcag 489
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 312 CATCAGCCCAACAACTGGATTACAAATCTGCCAAACACAGCCTGGCCTTCATTCTAGCAG 371
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 490 atgctggttttgcgtggtggggaacagcaggggaaacgcctggtctcgaaacaca 549
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 372 ATGCTGGCTATGATGTGTGGCTGGGGACAGTCGAGGGAATACATGGTCCCGGAAAAATG 431
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 550 agacactctccatagaccagaatgagttctgggttccagttatgatgatgagtgctaggt 609
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 432 TATACTATTACACAGACTCAGTTGAATCTGGCTTTTCAGCTTTGATGAATGGCTAAAT 491
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 610 ttgaccttctcagtgatataaactatttttgcaaaaacagggccaggaagaagatcctatt 669
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 492 ATGACCTTCCACCCACCATAGACTTTCATTTGTACAGAAACTGGACAAAGAGATACACT 551
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 670 atgtcggtctattcacagggccaccacatgggctttatttgcattttccaccatgccagagc 729
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 552 ATGTTGGTCACTCTCAGGGCCACCACATATCGGTTTATTTGCCCTTTCTACCAATCCTGCTC 611
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 730 tggctcagaataacaaatgtatttttgccttttagcaccatagccactgttaagcatgcaa 789
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 612 TGGCTAAAAAATCAAGAGGTTTTTATGATATAGCTCCAGTTGCTACTGTGAAGTATACAG 671
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 790 aaagcccccgggaccaaaattttgtgtgtccagatcatgatcatcaagggattgtttggca 849
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: August 24, 2001, 21:22:40
Job time: 5547 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:38:56 ; Search time 97.86 Seconds
(without alignments)
2454.893 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgcag.....gacgggtgtgagccgtattg 1269

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appli
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appli
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appli
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appli
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appli
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appli
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appli
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appli
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appli
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appli
11	46.4	3.7	43676	3 US-09-356-952-12	Sequence 12, Appli
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appli
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appli
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 154, App
15	36.6	2.9	3282	1 US-08-276-852-159	Sequence 169, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
17	36.6	2.9	3282	1 US-08-899-575-159	Sequence 169, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
19	36.6	2.9	3282	1 US-08-899-575-159	Sequence 169, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
21	36.6	2.9	3282	5 PCT-US95-08743-159	Sequence 169, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
23	36.6	2.9	13254	1 US-08-276-852-170	Sequence 170, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
25	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
27	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App

28	36.6	2.9	13254	5 PCT-US95-08743-156	Sequence 156, App
c 29	36.6	2.9	13254	5 PCT-US95-08743-170	Sequence 170, App
30	33.6	2.6	8392	1 US-08-080-255-6	Sequence 6, Appli
31	33.6	2.6	8392	3 US-08-465-713-6	Sequence 6, Appli
32	33.6	2.6	8392	5 PCT-US93-05857-6	Sequence 6, Appli
c 33	32.6	2.6	7218	1 US-08-232-453-14	Sequence 14, Appli
c 34	32	2.5	2197	1 US-08-233-005-3	Sequence 3, Appli
c 35	32	2.5	2197	1 US-08-428-943-3	Sequence 3, Appli
c 36	32	2.5	2197	3 US-09-016-649-3	Sequence 3, Appli
c 37	32	2.5	2197	5 PCT-US95-04858-3	Sequence 3, Appli
38	32	2.5	8342	3 US-08-545-860D-63	Sequence 63, Appli
39	32	2.5	8342	5 PCT-US94-04496-63	Sequence 63, Appli
40	31	2.4	3212	3 US-08-673-814-4	Sequence 1, Appli
c 41	30.4	2.4	1529	1 US-08-336-778-1	Sequence 1, Appli
c 42	30.2	2.4	4982	3 US-08-699-103B-1	Sequence 1, Appli
c 43	30.2	2.4	10079	2 US-08-476-866-20	Sequence 20, Appli
c 44	30	2.4	3639	2 US-08-737-524B-26	Sequence 26, Appli
45	30	2.4	5178	2 US-08-474-169-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-227-108-2
Sequence 2, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 accagaagcattcatgaataattagtgaaatcatcccaacatcaaggctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGGATACCCAGCTGAGG 85
QY 185 aatatgaagtcgaactgaagatgggtatatctctttctgttaacagagattcctcgagcc 244
DB 86 AATATGAAGTGTGACCGAAGAGCGGTATATCTCTGGGATCGACAGAAATTCCTTATGGGA 145
QY 245 tagtgaacctgaagacagaggtccaggcctgtgtgttactcagcatggcctagt 304
DB 146 GGAATAATTCAGAGATATAGCGGGAGACCTGTGTGATTTTGCAACAGGTTTGCTCG 205
QY 305 gaggtgtagcaactgatttccaaactgcccacatagcctgggtcttctattctgagcag 364
DB 206 CATCAGCCACAACCTGGATCTCCAACTGCCACACAGCTGGCTTCTATCTGGCCG 265
QY 365 atgctggtttgacgctgtggtatgggggagcaggggaaacagcgtgctgtcgaatacaca 424
DB 266 ACGCCGGGTACGACGTGTGGCTGGGAACAGCAGGGGCAACACCTGGCCAGGGAATC 325
QY 425 agacactctccatagaccagaagatgagttctgggtcttccagttatgatgagatggctagt 484
DB 326 TGTACTACTGCCCGACTCCGTGGAATCTGGGCTTTTCAGCTTTGACGAGATGGCTAAAT 385
QY 485 ttgacctctctcagtgatataaactttatttgcagaaacggggcaggaagaaatctatt 544
DB 386 ATGACCTTCCCGCCACCATTTGACTTCATCTTGAAGAAACGGGACAGACAGCTACT 445
QY 545 atgtcggtattcacagggcagccacacatgggtttatttgcatcttccaccatgccagagc 604
DB 446 ACGTTGGCCATTCACCGGACACCACTTGTGTTCATCGCTTTTCACCAATCCCAAGC 505
QY 605 tggctcagaaaatcaaatatttcttcttagcaccatagccactgttaagcatcaaa 664
DB 506 TGGGAAACGGATCAAACTCTTATGATAGTCTCCGTTGCCACCGTGAAGTACACCG 565
QY 665 aaagcccgaggaccataattttgtgtgcagatgatgataagggattgtttggca 724
DB 566 AAACCTCTTAAACAACTCATGCTCGCTCCCTGTTCCTTCAAGCTTATATTGGAA 625
QY 725 aaaaagaattctctgatacagaccagattctcagacaactgtttatttaccctgtgagcc 784
DB 626 ACAAATATTCTACCCACACCACTTCTTGTATCAATTTCTGCCACCGAGGTATGCTCC 685
QY 785 aggtgattctgtacagattgttagtaataatcatctgttactcttggtggattcaacca 844
DB 686 GCGAGCGGTGGATCTCTCTGACGAAACGCCCTGTGTATCATTTGTGGATTGACACTA 745
QY 845 acaatatgaacatgagccgagcaagtgatgctgtccacactctgtctgggaacatctg 904
DB 746 TGAACCTTGAACATGAGTCGCTTGGATGTGTATCTGCATATAATCCAGCAGGAACATCGG 805
QY 905 tgcataatctctacactgagccagcagtgaaatctgtgtaactccggcatttgact 964
DB 806 TTCAAGACGTGCTCCCTGCTCCAGGCTGTAAAGTCTGGGAAGTTCCAAAGCTTTGACT 865
QY 965 ggggagtgagaccaaaactggaataatgcaatcagcaactccctgtaaggtacagag 1024
DB 866 GGGGAACCCGATTCAGAAATGATGACATATCATCAGAGATGCTCCCTCTACTACAAAC 925
QY 1025 tcagagatatgacggtccctacagcaatgtggacagaggttcagagactggctttcaaatc 1084
DB 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTGCTGGCCGAC 985
QY 1085 cagaagacgtgaaatgctctctctgaggtgaccaactctatcaccataagaatattc 1144
DB 986 CTCACGATGTGACTTTTGTGCTTCCAAAGCTCCCAATCTCATTTACCAAGGAATTC 1045
QY 1145 ctgaatgggctcagtgagattctctgggtttggatgctctcaccgtatgtacaagt 1204
DB 1046 CTCCTTCAATCACTTGGACTTATCTGGGCCATGGATGCCCCCTCAAGCGGTTTACAATG 1105

QY 1205 aatcatccatctgatg 1221
DB 1106 AATTTTCCATGATG 1122
RESULT 2
US-09-073-674-2
; Sequence 2, Application US/09073674
; Patent No. 598189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/POCKET NUMBER: 5072-DI-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2

Query Match 35.0%; Score 444.2; DB 2; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5,6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
QY 125 accagaagcattcatgaataattagtgaaatcatcccaacatcaaggctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGGATACCCAGCTGAGG 85
QY 185 aatatgaagtcgaactgaagatgggtatatctctttctgttaacagagattcctcgagcc 244
DB 86 AATATGAAGTGTGACCGAAGAGCGGTATATCTCTGGGATCGACAGAAATTCCTTATGGGA 145
QY 245 tagtgaacctgaagacagaggttccaggcctgtgtgttactcagcatggcctagt 304
DB 146 GGAATAATTCAGAGATATAGCGGGAGACCTGTGTGATTTTGCAACAGGTTTGCTCG 205
QY 305 gaggtgtagcaactgatttccaaactgcccacatagcctgggtcttctattctgagcag 364
DB 206 CATCAGCCACAACCTGGATCTCCAACTGCCACACAGCTGGCTTCTATCTGGCCG 265
QY 365 atgctggtttgacgctgtggtatgggggagcaggggaaacagcgtgctgtcgaatacaca 424

Db	217	TATTCTCTGCGCTTTAACCGAATPCTCTATGGGAGGAAGAACCATTTCTGTACAAAGGTGCCAA	276
Qy	273	gcctgtggtgttactgcagcatggcctagttggtgaggtgctagcaactgatttccaacct	332
Db	277	ACCAAGTGTCTTCTGCAACATGCGTTGCTGGCAGATTTCTAGTAACCTGGGTCACAAACTT	336
Qy	333	gcccacaaatagcctgggtctcatctctggcagatgctgtggttttgaactgtgtgattggtgaa	392
Db	337	TGCCAAGCAGCAGCGCTGGCGTTTCATTTCTGTGTGATGCTGGTTTGTGACGTGTGGATGGGCA	396
Qy	393	cagcaggggaaacgccttggtctgaaacacaaagacacactctccatagaccagaatgagtt	452
Db	397	CAGCAGAGGAATTAACCTGGTCTCGGAACATPAGACACTCTCAGTTTCTCAGGATGAAT	456
Qy	453	ctgggtcttcagtattatgatgatggtgtaggttttgacctctcctgcagtgataaaacttcat	512
Db	457	CTGGGCTTTCAGTTATGATGATGATGGCAAAATATGACCTACGAGCTTCCATTAACATTCAT	516
Qy	513	tttcgaaaaacggccaggaagaagactattatgtcggtctattcacagggcaccacacat	572
Db	517	TCTGAATTAATACTGGCCAAGAACAAAGTGTATTATGTGGGTCAATTTCTCAAGGCACCACTAT	576
Qy	573	gggctttattgcattttccacatgccagagctggctcagaaaaatcaaaatgtattttgc	632
Db	577	AGGTTTTATAGCATTTTACAGATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTTTGC	636
Qy	633	tttagcaccatagccactgttaagcatgcaaaaagccccgggagccaaaattttttgttgc	692
Db	637	CCTGGGTCCTGTGGCTCCGTCGCCCTCTGTACTAGGCCCTATGGGCCAAATTAGGACGATT	696
Qy	693	gccagatgatgatcaaggattgtttggcaaaaagaattctctgtatcagaccagatt	752
Db	697	ACCAGATCATCTCATTAGGACTTATTTGGAGACAAGAATTTCTTCCCAGAGTGGCT	756
Qy	753	tctcagacaactgtttattacctttgtggccaggtgattctctgtacagattgttagtaa	812
Db	757	TTTGAAGTGGCTGGGTACCCAGTTTGGCATCTCATCTACATCTGAAGGAGCTCTCTGTGGAA	816
Qy	813	tatcatgttactctgtgtgattccaaccaacaatatgaacatgacgagccagcaagt	872
Db	817	TCTCTGTTTTCTTCTGTGTGGATTTAATGAGAGAAATTTAAATATGCTCTAGAGTGGATGT	876
Qy	873	atatgtgccccacactcttctgtgaacatctgtgcaaaatattctacactggagccagc	932
Db	877	ATATACAACACATTTCTCTGCTGGAATTCCTGTGCAAAACATGTTACACTGGAGCCAGC	936
Qy	933	agtgaaattctgtgaactccgggacatttgactggtggggagtgagacaaaactctggaaaa	992
Db	937	TGTTAAATTTCCAAAAGTTTCAAGCCTTTTGACTTGGGAAGCAGTGGCCAGAATTTATTTCA	996
Qy	993	atgcatacagccaactcctgtaaggtacagagtcagagatatgacggtccctcacgcaat	1052
Db	997	TTACAACACAGATTATCTCCACATACATAATGTGAGGACATGCTTGTCCGACATGCAGT	1056
Qy	1053	gtggacaggaggtcaggactggtcttcaaatccagaagacgtgaaaaatgctgctctctga	1112
Db	1057	CTGGAGCGGGGTCAGCAGTGGCTTCAGATGCTACGACGCTCAATATCTTACTGACTCA	1116
Qy	1113	gttgaccaacctcatctaccataagaatatctgtaatgggtccagctgagatttcctctg	1172
Db	1117	GATCACCAACTTGGTGTCCATGAGAGCAATTCGGAAATGGGAGCATCTTGTGACTTTCATTG	1176
Qy	1173	gggttttgatgctctccacgtatgtacaaatgaaatccatccatctgatgcagcaggagga	1232
Db	1177	GGGCTGGATGCCCTTGAGGCGCTTTATATAAATAATTAATCTAATGAGGAATATCA	1236
Qy	1233	gaccaacct	1241
Db	1237	GTGAAGCT	1245

RESULT
A26689
5

Mon Aug